

United States Department of Commerce

Bureau of the Census

REPORT TO CONGRESS --

THE PLAN FOR CENSUS 2000

Originally Issued  
July 1997

Revised and Reissued  
August 1997

# TABLE OF CONTENTS

<b>Executive Summary</b>	ix
<b>I. The Importance of an Accurate Census</b>	1
A. A Tradition of Innovation in the Census	1
B. Lessons From the 1990 Census	2
1. Overview	2
2. Some Groups Counted Less Effectively Than Others	3
3. Effects of Inaccuracy	4
4. An Exhaustive Attempt to Make Traditional Methods Work	4
5. Resources Were Adequate	5
6. Causes of Inaccuracy and the Undercount	5
C. A Consensus to Improve Census Accuracy	6
D. Recommendations of National Academy of Sciences' Panels	7
<b>II. The Census 2000 Plan</b>	9
A. Public and Congressional Involvement in Census 2000 Planning	9
1. The Advisory Committees	9
2. Public Meetings	9
3. Congressional Partnership	10
B. Major Components of the Plan	10
1. The Master Address File	10
2. Public Outreach and Marketing	10
3. Questionnaire Mail-out/Mail-back	11
4. Collecting Data on Populations Living in Nontraditional Households	12
5. Collecting Long Form Data to Meet Federal Requirements	13
6. Retrieving and Processing the Data from the Returned Forms	13
7. Matching and Unduplication	13
8. Processing	13
9. Quality Assurance	14
10. The Census 2000 Dress Rehearsal in 1998	14
11. Data Dissemination through DADS	14
12. Evaluation and Preparation for 2010	15

<b>III. Summary of Costs and Improvements</b>	<b>16</b>
<b>IV. Improvements of Traditional Methods in Census 2000</b>	<b>19</b>
A. The Master Address File	19
B. New Outreach Methods	21
C. New Technology	22
<b>V. Use of Scientific Sampling to Increase Accuracy</b>	<b>23</b>
A. Reliance on Sampling in Previous Censuses	23
B. Support for Sampling within the Scientific Community	24
C. Sampling to Collect Long Form Data	25
D. Sampling in the Postal Vacancy Check	26
E. Sampling for Nonresponse Follow-up	26
F. Integrated Coverage Measurement	29
<b>VI. Options for Addressing Areas with High Undercount Rates in the Absence of Scientific Sampling</b>	<b>33</b>
A. The Census Bureau Could Face Severe Labor Difficulties if Forced to Take Census 2000 without Sampling	33
B. The Only Alternative to Sampling Would Be to Rely More Heavily on Traditional Methods, Methods with Proven Limits	34
C. Spending More on Outreach Instead of Sampling Would Leave an Unacceptably Large Undercount and Have Biased Results	37
D. Without Sampling, Costs Would Increase by at Least \$675 Million, and the Final Count Would Be Less Accurate than the 1990 Census	37
<b>VII. An Illustration: The Milwaukee Complete Count Campaign</b>	<b>40</b>

<b>VIII. Expected Error Rates</b>	<b>41</b>
A. Nonsampling Error	41
1. Coverage Error	41
2. Nonresponse Error	41
3. Observational Error	41
4. Data Processing Error	42
B. Sampling Error	42
C. Error Related to Estimation	42
1. Model Error	42
2. Matching Error	42
3. Contamination Error	43
D. Gross Error Versus Net Error	43
E. Summary of Estimated Error, by Geographic Level Down to the Census Tract, for Plan Alternatives	44
F. Error Rates at the Block Level	46
<b>IX. Procedures to Ensure Unbiased Statistical Decisions</b>	<b>49</b>
<b>X. Legal Considerations</b>	<b>52</b>
<b>Glossary of Terms</b>	<b>55</b>

**This Report was required by Title VIII of P.L. 105-18.  
For ease of reference, readers can find the following  
specifically-requested items in the following locations:**

<u>Item</u>	<u>Report Section</u>
“a list of all statistical methodologies that may be used in conducting the Census”	II(B); IV; V
“an explanation of these statistical methodologies”	II(B); IV; V
“a list of statistical errors which may occur as a result of the use of each statistical methodology”	VIII
“the estimated error rate down to the census tract level”	VIII(E)
“cost estimation showing cost allocations for each census activity planned”	III
“an analysis of all available options for counting hard-to-enumerate individuals, without utilizing sampling or any other statistical methodology, including efforts like the Milwaukee Complete Count Project”	VI; VII
“an estimate of the error rate at the census block level based upon the 1995 test data”	VIII(F)

## **EXECUTIVE SUMMARY**

This "Report to Congress -- The Plan for Census 2000" responds to Congress' direction that the Department of Commerce provide it with a comprehensive and detailed plan for Census 2000, including methodologies, types and size of statistical errors, cost estimates, options for counting hard-to-enumerate individuals without statistical sampling, and error rates at the census block level in the 1995 test.

The Census Bureau's goal in Census 2000 is to take the most accurate and cost-effective census possible. The importance of an accurate decennial census cannot be overstated. Census data are used to reapportion the House of Representatives, ensuring that political representation is distributed evenly to all Americans, and to determine allocation formulas for the distribution of billions of dollars of federal and state funds each year. Census data tell us what we know about our country; they are the definitive benchmark for virtually all demographic information used by educators, policy makers, journalists, and community and nonprofit organizations.

The national census count became more accurate with each consecutive census from 1940 to 1980. Although it surpassed all previous censuses in terms of design, execution and resources used, the 1990 census took a large step backwards in terms of accuracy. While the 1980 census had fallen 2.8 million people below an accurate count, the census count in 1990 fell 4.7 million people short, missing 1.8 percent of the population, according to demographic analysis estimates. Moreover, the undercount in 1990 was not spread evenly across the nation; children and minorities were disproportionately undercounted.

In the wake of the 1990 census, there was a consensus among the Census Bureau, professional statisticians, and Congress that significant changes were required for the upcoming 2000 census; the Census Bureau could not continue to employ the methods it had been using. In 1991, bipartisan legislation passed unanimously by Congress and signed by President Bush directed the National Academy of Sciences (the Academy) to study "the means by which the Government could achieve the most accurate population count possible."

Changes in American society dictate that census-taking methods must change. The willingness of many Americans to respond to the decennial census has declined in recent years. Populations with high undercount rates under traditional methods of enumeration have grown more rapidly than the total population. The necessity of adapting census methodologies in response to societal changes was discussed by the Academy Panel to Evaluate Alternative Census Methodologies in its second interim report in June 1997: "[c]hange is not the enemy of an accurate and useful census; rather, not changing methods as the United States changes would inevitably result in a seriously degraded census."

Census 2000 will incorporate many changes. It will incorporate a Master Address File that is more comprehensive than ever. It will use a public outreach

and marketing campaign to educate people about the importance of being counted. It will utilize a more effective questionnaire mail-out and mail-back campaign -- the foundation of the census. It will incorporate advanced technologies to increase accuracy and speed. It will utilize statistical sampling to account for those who cannot otherwise be accounted for. And it will incorporate quality assurance plans to ensure an accurate one-number census. The details of the plan for Census 2000 are contained in this Report.

The Academy recommended that the Census Bureau consider "the appropriateness of using sampling methods in combination with basic data-collection techniques." As detailed in this Report, several Academy panels have examined the census process over the past six years and all have concluded that an accurate and cost-effective census cannot be taken without the introduction of a limited use of sampling. The Academy Panel on Census Requirements in the Year 2000 and Beyond concluded that,

[i]t is fruitless to continue trying to count every last person with traditional census methods of physical enumeration. Simply providing additional funds to enable the Census Bureau to carry out the 2000 census using traditional methods, as it has in previous censuses, will not lead to improved coverage or data quality.

The Plan for Census 2000 results directly from the 1991 legislation and the subsequent guidelines and recommendations of the Academy. All significant departures from the methodologies used in previous censuses have been endorsed by the Academy, the Bureau's advisory committees, and the scientific community.

The Plan for Census 2000 has received strong support from professional statisticians and demographers -- experts are convinced that the introduction of a limited use of scientific sampling in Census 2000 will result in a more accurate, less costly census. These experts also believe that the use of sampling in Census 2000 should minimize the opportunity for political manipulation, not increase it. Scientists understand that sampling has known, objective properties that are preferable to the certainty of missing several million individuals using traditional enumeration methods alone. They understand that uncontrolled error is more of a concern with a traditional headcount than with sampling.

Outreach efforts, like the City of Milwaukee's Complete Count Campaign in 1990, increase awareness and mail response rates, but do not solve the census accuracy problem. The alternative to introducing a limited use of sampling is to continue with traditional physical enumeration methods. Taking Census 2000 the same way that the 1990 census was taken would result in an expected undercount of at least 1.9 percent of the population (more than 5 million people), and would cost at least \$675 million more than the current plan. Increased outreach will not solve the problem; spending more money for less accuracy is not a feasible alternative.

To further ensure accuracy and to avoid even the appearance of possible manipulation of the census, the Census Bureau has made plans for Census 2000 to incorporate an unprecedented expert review process. The Census Bureau has proposed that the Academy convene another expert panel to guide the Bureau's work through the completion of Census 2000. This new group will review critically the statistical procedures for Census 2000, especially the use of statistical sampling. The new group will work closely with the statisticians and demographers at the Census Bureau through the entire census operation. Census 2000 will be conducted in the open, in full public view.

As requested, this Report details expected error rates for Census 2000 down to the census tract level. If Census 2000 is conducted using a traditional enumeration, without the introduction of sampling, the Census Bureau expects an average error rate of at least 1.9 percent at all levels of geography from the census tract level up to the national level. In contrast, the Census Bureau's plan for Census 2000, a plan involving the introduction of a limited use of sampling, has the following expected average error rates: 0.1 percent at the national level, 0.5 percent at the state level, 0.6 percent at the Congressional district level, and 1.1 percent at the census tract level.

The Report also discusses error rates at the census block level. The 1995 Census Test did not provide meaningful error rates at the block level. The block error rates measured in the 1995 Census Test reflect two facts, that the test had lower rates of sampling than will be used in Census 2000, and that some blocks had few people. Even traditional methods of enumeration have seemingly high block level error rates: the 1990 Census had an average block error rate of almost eight percent. Fortunately, with or without sampling, such substantial error at the census block level does not mean substantial inaccuracy when blocks are aggregated. At all geographic levels important to political representation and funds allocation, Census 2000 will provide more accurate results than physical enumeration alone.

The Census Bureau is confident that its plan for Census 2000 satisfies both the Census Act and the Constitution. The Department of Justice (DOJ) in 1994 specifically approved the Census Bureau's plan to use sampling in Census 2000, a position consistent with earlier DOJ opinions in prior administrations, both Democratic and Republican. Courts have interpreted the Constitutional requirement of an "actual Enumeration" as a command to take the most accurate census feasible. Due to changes in American society, the most accurate census feasible can no longer be taken by traditional physical enumeration methods alone. The introduction of a limited use of sampling is necessary for an accurate and cost-effective census in 2000.



## **I. THE IMPORTANCE OF AN ACCURATE CENSUS**

The Constitution commands that a census of the nation's population be taken every ten years. The constitutional purpose for the census is the apportionment of seats in Congress, but the information collected every ten years provides more than just state-by-state population totals. The decennial census provides the cornerstone of knowledge about the people of our nation.

State and local governments use census data to draw legislative districts of equal population to comply with the constitutional "one man, one vote" mandate, and to comply with the statutory requirements of the Voting Rights Act. Each year the federal government distributes billions of dollars in grants according to population-based formulas based on census data. Federal, tribal, state and local officials study the patterns of detailed census data before constructing hospitals, highways, bridges and schools. Private businesses large and small have come to depend on the Census Bureau's population, income, education and housing data to locate new offices, shops and factories, and to market new products. Census data also serve as definitive benchmarks for virtually every household survey by educators, policy-making agencies, and community and nonprofit organizations.

### **A. A Tradition of Innovation in the Census**

The first census was taken in 1790 by United States Marshals, who were directed to visit every dwelling place and count the individuals living there. Since that time, the census has evolved to accommodate societal changes and to use advances in technology and statistics.

- The nation's population grew faster than the number of U.S. Marshals. Over time, professional enumerators supplemented the work of Marshals and completely replaced them by the end of the 19th century.
- After spending eight years tallying by hand the results of the 1880 Census, Census Bureau employees invented the punch card machine.
- In 1940, the Census Bureau introduced its "short form" questionnaire for the majority of the population, using the "long form" set of questions for only a sample of the population. Prior censuses had asked all residents to answer all questions.
- In 1970, the Census Bureau introduced self enumeration by mail. Rather than send an enumerator to every household in the nation, the Census Bureau first mailed questionnaires for households to fill out and mail back, and then sent enumerators only to nonresponding addresses.

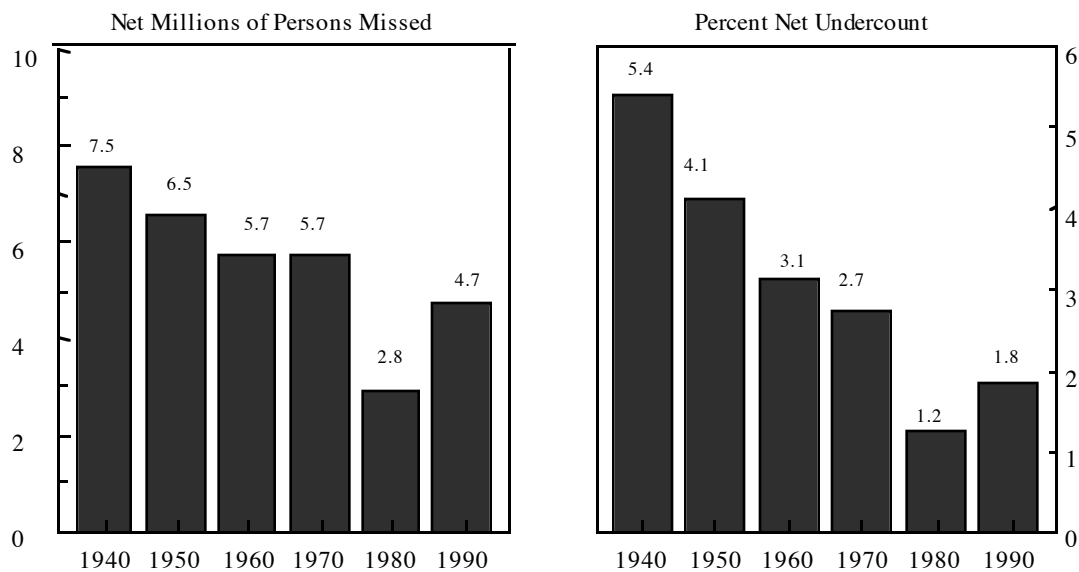
The basic structure for census data collection (mail-out/mail-back followed by enumerator visits to all nonresponding addresses) did not change from 1970

through 1990. The results of the 1990 Census, however, demonstrated that new methods are required.

## B. Lessons From the 1990 Census

**1. Overview.** The 1990 Census was a difficult undertaking. Though better designed and executed than any previous census, the Census in 1990 took a step backward on the fundamental issue of accuracy. For the first time since the Census Bureau began conducting post-census evaluations in 1940, the decennial census was *less* accurate than its predecessor. In spite of unprecedented efforts to count everyone, accuracy in the 1990 Census fell short of the accuracy achieved in the 1980 Census. On the basis of "Demographic Analysis,"<sup>1</sup> the undercount was 4.7 million people; the undercount rate of 1.8 percent in 1990 was 50 percent greater than the rate had been in 1980.

### Net Undercount Based on Demographic Analysis, 1940 - 1990

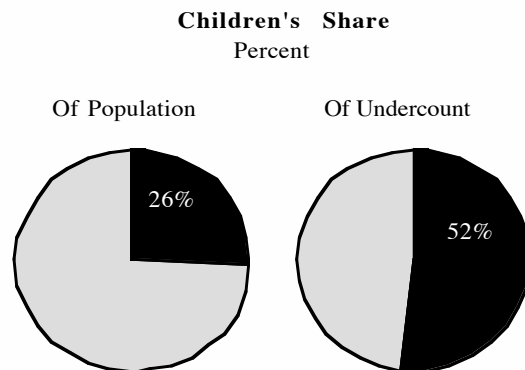


Source: J.G. Robinson, B. Ahmed et al.

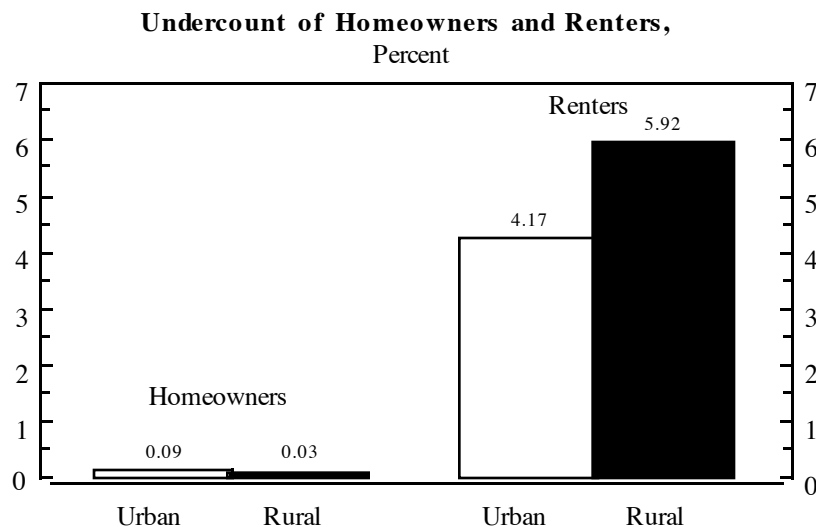
<sup>1</sup> Demographic Analysis is one of the two standard methods that the Census Bureau uses to measure coverage, that is the extent that the official census totals cover or completely account for the true total. Demographic Analysis relies on administrative records of births, deaths, immigration, and emigration to provide estimates of the true total. Demographic Analysis is the only method for analyzing historical trends in the shortfall in coverage, the national undercount.

**2. Some Groups Counted Less Effectively Than Others.** To measure the extent to which identifiable groups were not fully counted, the Census Bureau conducted a Post Enumeration Survey (PES)<sup>2</sup> as part of the 1990 Census process. The PES found large differences in the undercount rates for different groups, a phenomenon called the “differential undercount.”

***Children were much more likely than adults to be undercounted in the 1990 Census.*** While children under the age of 18 represented 26 percent of the total national population that year, they accounted for 52 percent of the undercount.

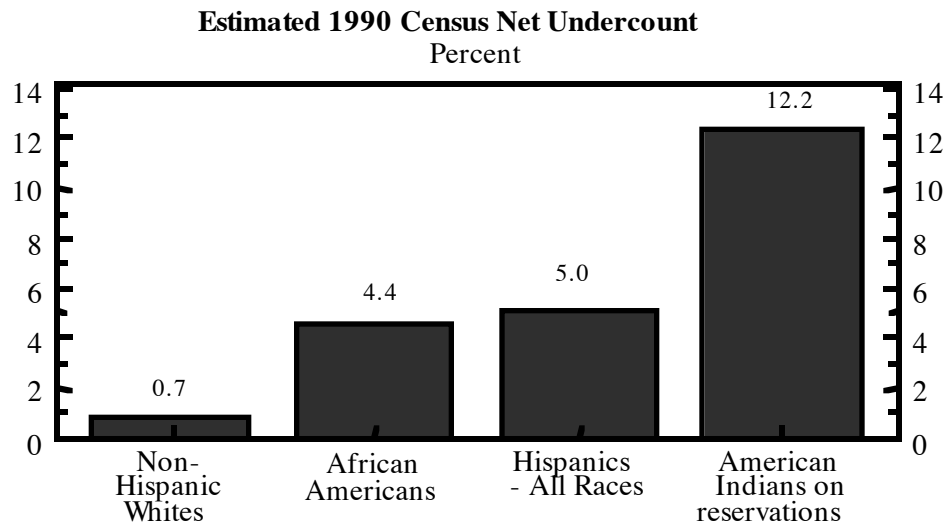


***Renters, particularly in rural areas, were also more likely to be left out of the official Census count in 1990.*** The PES found a 5.9 percent undercount among renters in rural areas.



<sup>2</sup> The PES evaluated coverage on a case-by-case basis using the Dual System Estimation methodology explained in Section V(F). The PES provided undercount information for detailed categories, such as renter/homeowner and racial and ethnic group, that are not possible with demographic analysis.

***Racial and ethnic minorities were also affected disproportionately.*** Compared to the undercount rate for non-Hispanic Whites, the 1990 undercount rates were six times larger among African Americans and seven times larger among Hispanics. Twelve percent, or nearly one out of every eight, American Indians living on reservations were not counted in 1990.

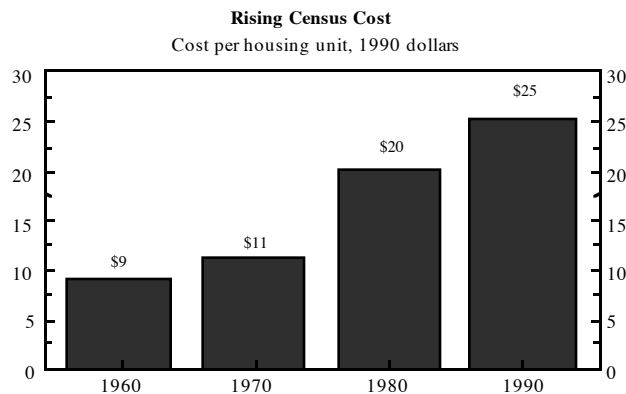


**3. Effects of Inaccuracy.** As a result of the inaccuracy in the 1990 Census, many Americans were denied an equal voice in their government. Federal spending employing population-based formulas -- for schools, crime prevention, health care, and transportation -- was misdirected.

**4. An Exhaustive Attempt to Make Traditional Methods Work.** The 1990 Census failed to match the accuracy of the 1980 Census despite the Census Bureau's exhaustive attempt to make traditional methods work. The Census Bureau deployed more than a half million people around the country to collect information from the approximately 36 million addresses that had not responded by mail. The Bureau devoted considerable effort and resources on operations to improve the count in areas with the greatest risk of a large undercount:

- Four advisory committees helped develop and implement specialized outreach efforts to racial and ethnic minority groups.
- To promote awareness about the census and its importance, the Census Bureau invested \$75 million in promotion and outreach activities, and worked with the Advertising Council on a public service announcement campaign valued at \$68 million.
- Toll free telephone numbers were created so that people who had questions about how to fill out the forms could get assistance or request a Spanish-language form.

**5. Resources Were Adequate.** The deterioration in accuracy of the census from 1980 to 1990 cannot be attributed to inadequate funding by Congress. The Census Bureau requested, and received, additional appropriations from Congress. The 1990 Census was the most expensive in history, costing \$25 per housing unit. On an inflation adjusted basis, the 1970 Census had cost only \$11 per housing unit, and the 1980 Census \$20 per housing unit.



Source: B. Edmonston, and C. Schultze, Editors, *Modernizing the U.S. Census*, 1995.

**6. Causes of Inaccuracy and the Undercount.** How did such a comprehensive effort result in the first count known to be *less* accurate than its predecessor? Experts at the Census Bureau and three National Academy of Sciences (Academy) panels commissioned to study the problem concluded that the lower accuracy and higher costs of the 1990 Census were the product of several troubling societal trends:

- An increasing number of Americans were too busy to be counted. The number of people working more than one job had increased, along with the number of multiple-worker families, so people were home less often when enumerators visited. When people were home, they were less willing to spend time filling out a census form.
- Americans were inundated with junk mail, mail that obscures important documents such as census forms.
- More Americans lived in housing that was remote or inaccessible. For example, security guards in gated communities did not always cooperate with enumerators.
- More Americans were becoming alienated from society in general and more mistrustful of government in particular. They had also grown more concerned about privacy.

These experts also concluded that the population of people more likely to be left uncounted has been growing more rapidly than the total population. Census

tracts with high undercount rates tend to have the following characteristics to a greater degree than the rest of the country:

- A highly mobile population.
- Language barriers.
- High concentrations of unmarried residents.
- Nontraditional housing arrangements, such as extended families, cohabiting couples, roommates, boarders and other nonrelatives.
- Irregular housing, such as illegal units, mobile homes and secured buildings.
- Neighborhood conditions that lead to resistance to outsiders, concealment to protect resources, and disbelief of census confidentiality.

Because higher proportions of the nation's children, renters, and minorities live in these situations, it should not be surprising that their undercount rates are higher.

The sharp decline in the rate that people return their census questionnaires presents a clear example of how the changes in society directly affect the operation of the census. When census questionnaires were first mailed in 1970, 78 percent of housing units mailed back their questionnaires. By 1990, that percentage had fallen to 65 percent.

Every indication since 1990 suggests that the census-taking environment is likely to be even more difficult in 2000 than it was in 1990. For example, the percentage of married couple families with both spouses employed rose steadily from 28 percent in the 1960 Census to 50 percent in the 1990 Census. That trend has continued upward in the 1990s.

### **C. A Consensus to Improve Census Accuracy**

Congress, concerned about the accuracy and cost problems of the 1990 Census, passed the Decennial Census Improvement Act of 1991 (P.L. 102-135), signed by President Bush, requiring the National Academy of Sciences to study "the means by which the Government could achieve the most accurate population count possible," specifically considering, *inter alia*, "the appropriateness of using sampling methods in combination with basic data-collection techniques or otherwise, in the acquisition or refinement of population data, including a review of the accuracy of the data for different levels of geography ..." The legislation enjoyed bipartisan support: the House of Representatives passed it unanimously and the Senate passed it under a suspension of the rules by unanimous consent.

#### **D. Recommendations of National Academy of Sciences' Panels**

Since 1990, the Academy's Committee on National Statistics has established three separate panels to study how to improve the next decennial census: the Panel on Census Requirements in the Year 2000 and Beyond ("Academy Panel on Requirements"), the Panel to Evaluate Alternative Census Methods ("Academy Panel on Methods"), and the Academy Panel to Evaluate Alternative Census Methodologies ("Academy Panel on Alternative Methodologies").<sup>3</sup>

The Academy Panel on Requirements found that traditional methods were incapable of eliminating the undercount:

It is fruitless to continue trying to count every last person with traditional Census methods of physical enumeration. Simply providing additional funds to enable the Census Bureau to carry out the 2000 Census using traditional methods, as it has in previous Censuses, will not lead to improved coverage or data quality.

It also found that more radical alternatives to a traditional enumeration (a national register, an administrative records census, a census conducted by the U.S. Postal Service, and a rolling sample census) were either not feasible or not consistent with American values.

Finally, the Academy Panel on Requirements found that scientific sampling, both for nonresponse follow-up and to improve accuracy (each is described in Section V) would both increase accuracy and lower costs. That panel concluded that scientific sampling was not just *a* solution to the cost and accuracy problems, it was the *only* solution.

---

<sup>3</sup> The Academy Panel on Requirements was tasked with considering the purposes of a decennial census and alternative data collection systems. This panel supplemented an existing panel, the Academy Panel on Methods that had been tasked with studying how the census should be taken and that focused on more specific methodologies. The Academy Panel on Requirements issued an interim report in May 1993 and its final report, "Modernizing the U.S. Census," in 1995. The Academy Panel on Methods issued a "Letter" report in December 1992, an interim report in September 1993, and its final report, "Counting People in the Information Age," in the fall of 1994. Subsequently, the Academy created its Academy Panel on Alternative Methodologies, which has issued two interim reports: "Sampling in the 2000 Census: Interim Report I," in June 1996, and "Preparing for the 2000 Census: Interim Report II," in June 1997.

The report from the Academy Panel on Methods concurred with the Academy Panel on Requirements that statistical sampling should be used both for nonresponse follow-up and to increase accuracy:

Differential undercount cannot be reduced to acceptable levels at acceptable costs without the use of integrated coverage measurement and the statistical methods associated with it. We endorse the use of integrated coverage measurement as an essential part of Census-taking in the 2000 Census ... Sampling for nonresponse follow-up could produce major cost savings in 2000. The Census Bureau should test nonresponse follow-up sampling in 1995 ...

The Second Interim Report from the Academy Panel on Alternative Census Methodologies likewise concluded that census methods need to change in response to societal changes:

Changing, updating, and adapting the Census methods is a proven and desirable course of action. Change is not the enemy of an accurate and useful Census; rather, not changing methods as the United States changes would inevitably result in a seriously degraded Census.

Thus, the current plan for Census 2000 results from Congressional legislation, enacted in the aftermath of the 1990 Census, that directed the Census Bureau to achieve greater accuracy by revising census methodology in accordance with recommendations formulated by the National Academy of Sciences.



## II. THE CENSUS 2000 PLAN

Census 2000 will be the largest peacetime mobilization in our nation's history. Its goal is to ensure that every individual in the United States on April 1, 2000 is accounted for.

### A. Public and Congressional Involvement in Census 2000 Planning

**1. The Advisory Committees.** To ensure public involvement in the plan for Census 2000, the Bureau chartered a number of advisory committees under the Federal Advisory Committee Act, 5 U.S.C. § App. I, to assist in planning.

- **The 2000 Census Advisory Committee** consists of more than 30 professional, governmental, and nongovernmental organizations, such as the American Sociological Association, the Business Roundtable, the National Association of Counties, the National Governors' Association, the U.S. Chamber of Commerce, and the National Coalition for the Homeless. All Congressional oversight committees and subcommittees have had *ex officio* members on this Advisory Committee. The Committee advises the Secretary of Commerce about designing Census 2000 methodologies and procedures, reducing the differentials among population groups and geographic areas, and containing costs.
- **The Census Advisory Committee of Professional Associations** consists of the American Statistical Association, the Population Association of America, the American Economic Association, and the American Marketing Association. The Committee advises the Census Bureau on the full range of Census Bureau programs and activities in relation to its areas of expertise.
- **Four Census Advisory Committees on particular racial and ethnic populations** have been created: the **Census Advisory Committee on the African American Population**; the **Census Advisory Committee on the American Indian and Alaska Native Populations**; the **Census Advisory Committee on the Asian and Pacific Islander Populations**; and the **Census Advisory Committee on the Hispanic Population**. These four Committees provide the Census Bureau with an organized and continuing channel of communication with the communities they represent. They have helped the Census Bureau refine plans for nonresponse follow-up sampling, the paid advertising campaign, community-based outreach programs, and other matters critical to reaching all segments of the nation's population.

**2. Public Meetings.** To ensure even wider participation in the planning process, the Census Bureau has convened public meetings in thirty one cities across the country over the past three years to solicit input on the plan for Census 2000.

**3. Congressional Partnership.** In recognition of the Congress' particular interest in the decennial census, the Department of Commerce and the Census Bureau have strived to keep Congress informed about plans for improving Census 2000. Since the inception of the Secretary's 2000 Census Advisory Committee, the Chairman and ranking member of the House and Senate authorizing and appropriations committees and subcommittees have served as *ex officio* members of the Committee. Census Bureau staff cooperated with the Congress to develop P.L. 103-430, the Census Address List Improvement Act of 1994, that will enable effective partnerships with local and tribal governments, as well as with the U.S. Postal Service (USPS).

Since 1991, Department of Commerce officials have accepted 18 invitations to testify before House and Senate authorizing committees. In addition, officials have accepted invitations to testify before the House Appropriations Committee and the Senate Appropriations Committee. In the past two years, officials have responded to nearly 100 Congressional letters and more than 75 telephone requests inquiring about the plan for Census 2000. During that same period, the Census Bureau's Director and other Department of Commerce officials have held more than 100 briefings for Members and their staff on the plan for Census 2000.

## **B. Major Components of the Plan**

The Census Bureau's detailed plan for Census 2000 is set forth in "The Census 2000 Operational Plan", a copy of which is appended to this report. The highlights of the plan are explained below and in Sections IV and V.

**1. The Master Address File.** To conduct the census, the Census Bureau must identify and locate an estimated 118 million housing units in the nation. It plans to accomplish this goal by developing and maintaining a Master Address File (MAF) that is more comprehensive than ever before. For more detail on the MAF, see Section IV(A).

**2. Public Outreach and Marketing.** In 1990, the Bureau spent approximately \$75 million on promotion and outreach. It also supported a Public Service Announcement (PSA) effort that resulted in the airing of approximately \$68 million worth of donated advertisements. In spite of these efforts, the mail response rate dropped 10 percentage points and, according to Demographic Analysis, the net undercount was almost 2 million people higher than in the 1980 Census.

Part of this drop was caused by the Census Bureau's inability to ensure that PSAs were broadcast at optimum times and in appropriate markets. An evaluation of the 1990 PSA campaign by an outside contractor noted that ads were seldom placed at optimal times because decisions about when to air PSAs rested with local radio or television stations. Sixty percent of the U.S. population received 91 percent of the census advertising impact; forty percent received only 9 percent.

Post-census analysis found that outreach and publicity appeared to improve response and seemed as successful or more successful in 1990 than in 1980.

However, the 1990 Outreach Evaluation Survey also revealed that outreach was less successful among Blacks than among non-Hispanic Whites or Hispanics.

Based on its study of prior outreach campaigns, the Bureau concluded that the professional control of a paid media campaign would produce the best results. As former Director of the Census Bureau, Dr. Barbara Bryant (1989-1993) noted, “[t]he time for dependence upon pro bono creative work on Public Service Announcements for air time was past by 1990. The Census Bureau depended upon these for one census too many.”

Census 2000 will launch a \$208 million public outreach campaign to educate everyone about the importance of being counted. Among the improvements in public outreach and marketing are:

Targeted community outreach. In 2000, the Census Bureau will build partnerships with local and tribal governments, businesses, and community groups to get the word out, to endorse the census, and to encourage constituents to respond. Beginning in 1996 and expanding in 1998, government and community specialists will be hired to build relationships with local community and service-based organizations -- focusing on groups representing traditionally undercounted populations. An extensive outreach program is planned to reach schools, public sector employees, American Indians, and religious organizations. Businesses, nonprofit groups, and labor organizations will also be asked to endorse participation and to publicize the census through employee newsletters, inserts with paychecks, and through communications with members and local chapters.

Direct mail. The census questionnaire and related materials delivered to individual addresses will carry the same themes and messages as the overall campaign.

Public relations. The Bureau will use public meetings and the news media to inform the public about the value of the census and to encourage response. Communications specialists will be assigned to each field office to perform media outreach, to respond to media inquiries, and to coordinate the dissemination of Census 2000 messages. Local broadcaster/news director committees will be established in many communities to emphasize Census 2000 to TV viewers and radio listeners through broadcast segments and editorials in newspapers.

Advertising. The Census Bureau plans a targeted campaign to reach everyone through ads in newspapers, magazines, billboards, posters, radio, and television. A qualified private advertising firm will be retained to design and implement the Census 2000 advertising campaign at a projected cost of approximately \$100 million, primarily for buying media time.

**3. Questionnaire Mail-out/Mail-back.** In Census 2000, the questionnaire mail-out/mail-back system will be the primary means of census-taking, as it has been since 1970. The short form will be delivered to approximately 83 percent of all housing units. The short form asks only the basic population and housing

questions, while the long form will include additional questions on the characteristics of each person and of the housing unit. The long form will be delivered to a sample of approximately 17 percent of all housing units.

USPS letter carriers will deliver questionnaires to the vast majority of housing units that have city-style addresses (*e.g.*, 123 Main Street, Anytown, USA). In areas where there is no USPS delivery to city-style addresses, enumerators will hand-deliver addressed census questionnaires to each housing unit. In very remote or sparsely populated areas, enumerators will visit each housing unit and pick up or complete unaddressed questionnaires that the USPS previously delivered to each unit.

Because the American public is now deluged with junk mail, the Census Bureau has developed a questionnaire for 2000 that is easy to read, pleasing to look at, and simple to fill out. Private marketing experts are working with the Bureau to develop user-friendly designs that will help people understand why they are being asked for information. Sections IV(B) and IV(C) detail innovations in the data collection process.

#### **4. Collecting Data on Populations Living In Nontraditional Households.**

Taking a decennial census does not involve counting people just living in houses and apartments. In Census 2000, the Census Bureau must also enumerate people who live in group quarters and other nontraditional housing units, as well as people with no usual residence. These units include nursing homes, group homes, college dormitories, migrant and seasonal farm worker camps, military barracks or installations, American Indian reservations, and remote areas in Alaska.

Some of the enumeration methods that will be used for these special populations are:

- The Census Bureau has designed an operation for Census 2000 called Service-Based Enumeration (SBE) to improve the count of individuals who might not be included through standard enumeration methods. The SBE operation will be conducted in selected service locations, such as shelters and soup kitchens, and at targeted outdoor locations.
- Another special operation will count highly transient individuals living at recreational vehicle campgrounds and parks, commercial or public campgrounds, marinas, and even workers' quarters at fairs and carnivals.
- The Census Bureau is working with tribal officials to select the appropriate data collection methodologies for American Indian reservations.
- Remote areas of Alaska, often accessible only by small airplanes, snowmobiles, four-wheel-drive vehicles, or dogsleds, will be enumerated beginning in mid-February. This special timing will permit travel to these areas while conditions are most favorable.

- The Bureau will work with the Department of Defense and the U.S. Coast Guard to count individuals living on military installations, and with the U.S. Maritime Administration to identify maritime vessels for enumeration.

**5. Collecting Long Form Data to Meet Federal Requirements.** The census is the only data gathering effort that collects the same information from enough people to get comparable data for every geographic area in the United States. The long form has been used on a sample basis since 1940 to collect more data, more rapidly, while reducing overall respondent burden. In 2000, the long form will ask questions addressing the same seven subjects that appear on the short form, plus an additional 27 subjects which are either specifically required by law to be included in the census or are required in order to implement other federal programs and the census is the only source of the data.

**6. Retrieving and Processing the Data from the Returned Forms.** The Census Bureau has contracted with the private sector to secure the best available data capture technology. This technology will allow the Census Bureau to control, manage and process Census 2000 data more efficiently.

The Census 2000 data processing system will be a complex network of operational controls and processing routines intended to store and service the decennial control and data requirements. The Bureau will record a full electronic image of every questionnaire; sort mail-return questionnaires automatically; use optical mark recognition for all check-box items; and use intelligent character recognition to capture write-in character-based data items. The system will allow the Bureau to reduce the logistical burdens associated with handling large volumes of paper questionnaires. Once forms are checked in, prepared and scanned, all subsequent operations will be accomplished using the electronic image and captured data.

**7. Matching and Unduplication.** Census 2000 will make it simpler for people to be counted by providing them with multiple opportunities, and multiple methods, to respond. These response options will make it easier for everyone to be counted, but will increase the possibility of multiple responses for a given person and/or household. Unduplication of multiple responses in past censuses would have required a massive clerical operation, since only a small subset of person names was data captured. Advances in computer technology in the areas of computer storage, retrieval, and matching, along with image capture and recognition, have now given the Census Bureau the flexibility to provide multiple response options without incurring undue risk to the accuracy of the resulting census data.

**8. Processing.** The electronic images and data will be edited by computerized routines, checked for completeness and consistency, and prepared for tabulation and release of totals. As part of this process, missing information will be imputed.

**9. Quality Assurance.** To detect, correct, and minimize performance errors in critical census operations, the Census Bureau has developed individual quality assurance plans for all activities that could contribute to errors in outcome, such as misprinted census forms, inaccurate maps or address lists, faulty intelligent character recognition, inadequate training of enumerators, and miskeyed entries. The Census Bureau has created Quality Assurance Plans for each significant activity in Census 2000. In most cases, the Census Bureau will perform back-up checks for each procedure. Errors will be corrected, and steps will be taken to prevent similar errors in the future.

**10. The Census 2000 Dress Rehearsal in 1998.** A good dress rehearsal is crucial to a successful Census 2000, and the key to any dress rehearsal is making it as much like the actual event as possible. The Census Bureau has selected three sites in which to conduct the Census 2000 Dress Rehearsal: Sacramento, California; Columbia, South Carolina, along with 11 surrounding counties in north central South Carolina; and the Menominee American Indian Reservation in northeastern Wisconsin. The Census Bureau believes these three sites will provide a good operational demonstration of Census 2000 procedures and systems.

Since the summer of 1996, the Census Bureau has been working closely with local officials and community-based organizations in each of the three sites to plan and build the various infrastructures needed to ensure a successful dress rehearsal. These joint activities include refining the geographic database, building and refining the address list, and working with community and tribal organizations to plan effective outreach and promotion efforts. The Bureau has recruited staff in all three sites to begin complete address list development and verification.

The dress rehearsal will allow for a thorough demonstration of the most critical procedures for Census 2000. These procedures include address list development, marketing and promotion, and data collection, processing, and tabulation. The dress rehearsal plan will also demonstrate the use of statistical sampling in four major census operations: nonresponse follow-up, housing units designated as undeliverable as addressed by the USPS, Integrated Coverage Measurement (ICM), and the long form survey. (Discussed in more detail in Section V.)

**11. Data Dissemination through DADS.** The census provides a wealth of data that researchers, businesses, and government agencies are eager to research. Taking advantage of today's computer and Internet capabilities, the Census Bureau plans to make data from Census 2000 more readily available than any previous decennial census data. Census 2000 data will be tabulated and disseminated using the Data Access and Dissemination System (DADS). DADS will provide an interactive electronic system to allow data users to access prepackaged data products, documents, and on-line help, as well as to build custom data products on-line and off-line.

The Census Bureau has solicited the advice and recommendations of data users throughout the planning, design, and testing stages of DADS. DADS will be accessible to the widest possible array of users through the Internet and all available intermediaries, including the nearly 1,800 Data Centers and affiliates,

the 1,400 Federal Depository libraries and other libraries, universities, and private organizations. DADS will allow users to create customized products such as tables, charts, graphs and maps based on Census Bureau or user-defined geographic areas, and access metadata that provide documentation and explanatory information for data subjects and geographic areas.

**12. Evaluation and Preparation for 2010.** Once Census 2000 is completed, the Bureau will, as it has after all the censuses it has taken, conduct a variety of post-census evaluation studies. These studies will help data users, both within and outside the Census Bureau, to assess the data from Census 2000 and plan for the 2010 Census. In the past, these studies have relied on Demographic Analysis, statistical methods, and ethnographic analyses.

### III. SUMMARY OF COSTS AND IMPROVEMENTS

The following table provides the estimated costs of Census 2000, allocated among ten major activities and their components. The costs are expressed in constant dollars projected for year 2000.

CENSUS 2000 MAJOR ACTIVITIES: COST AND IMPROVEMENTS FROM 1990 CENSUS		
ACTIVITY	COST IN MILLIONS	IMPROVEMENTS FROM 1990 CENSUS
<p><b><u>Building the Address List</u></b></p> <ul style="list-style-type: none"> <li>- Plan and conduct address validation operations in areas with city-style addresses</li> <li>- Plan and conduct address listing operations in areas with high concentration of non-city-style addresses</li> <li>- Update and validate address list data</li> </ul>	<p><b>\$286</b></p> <p>\$110</p> <p>\$132</p> <p>\$44</p>	<p>For Census 2000, the 1990 address list is updated with the United States Postal Service list and local address lists to account for about 81 percent of all addresses.</p> <p>The census address list and the census geographic file are totally integrated for Census 2000, which will enable enumerators to locate housing units faster.</p> <p>All city-style addresses have corresponding streets in the census geographic file and non-city-style address locations are added to the census geographic file in Census 2000.</p> <p>Census address listers visit only a small portion of city-style addresses for field validation (for the 81 percent of addresses referenced above.)</p> <p>Local jurisdictions will be encouraged to review and correct the address list. Maintaining a nationwide, continuously updated and increasingly accurate census address list linked to the census geographic file is critical to any census.</p>
<p><b><u>Testing the Reengineered Methods</u></b></p> <ul style="list-style-type: none"> <li>- Questionnaire mailing tests <ul style="list-style-type: none"> <li>• 1995 Census Test</li> <li>• 1996 Community Census</li> <li>• National Content Survey</li> <li>• Race and Ethnic Targeted Test</li> </ul> </li> <li>- Conduct Census 2000 Dress Rehearsal</li> <li>- Sampling and estimation design</li> <li>- Research use of administrative records</li> </ul>	<p><b>\$218</b></p> <p>\$71</p> <p>\$37</p> <p>\$77</p> <p>\$33</p>	<p>Testing the new sampling and estimation methods will occur prior to being incorporated into Census 2000 on a larger scale.</p> <p>In addition to sampling for content (the "long form"), sampling and estimation will be used to 1) complete the nonresponse follow-up; and 2) produce a one-number census through Dual System Estimation.</p> <p>In Census 2000, the Bureau will use an automated matching system to ensure that each person is enumerated at his/her usual residence. This was largely a clerical operation in 1990.</p>



<p><b><u>Putting the Field Structure in Place</u></b></p> <ul style="list-style-type: none"> <li>- Field geographic support of address listing, data collection, and tabulation activities</li> <li>- Recruiting, training, and documentation preparation</li> <li>- Regional management of field offices, including <ul style="list-style-type: none"> <li>• space and support activities</li> <li>• equipment and telecommunications</li> </ul> </li> </ul>	<p><b>\$621</b></p> <p>\$67</p> <p>\$112</p> <p>\$171</p> <p>\$142</p> <p>\$129</p>	<p>The field office structure will be more streamlined and focused on data collection functions; for example, all questionnaires will be returned directly to the data capture centers.</p> <p>Data collection maps will be produced locally to better target enumerator activities.</p> <p>Development of an automated enterprise-wide integrated personnel and payroll system to administratively support the Census 2000 temporary workforce will permit efficient employment and payroll processing.</p> <p>Use of PC-based client-server architecture for the key automation systems will provide an effective, flexible processing capability.</p>
<p><b><u>Reaching the Public / Marketing</u></b></p> <ul style="list-style-type: none"> <li>- Design and operation of national advertising campaign</li> <li>- Partnership activities with state, local, and tribal governments and national &amp; umbrella organizations</li> </ul>	<p><b>\$208</b></p> <p>\$100</p> <p>\$108</p>	<p>In Census 2000, the Census Bureau will use paid advertising by the "best in class" contractor versus <i>pro bono</i> advertising used in 1990.</p> <p>The Census Bureau will have more targeted outreach efforts than in 1990 through the use of Government Specialists, Community Specialists, and Media Specialists.</p> <p>The Census Bureau will hire a "best in class" contractor to develop and distribute curriculum and promotional materials about Census 2000 to schools.</p>
<p><b><u>Printing and Mailing Questionnaires</u></b></p> <ul style="list-style-type: none"> <li>- Printing of initial and replacement questionnaires</li> <li>- Printing of advance notice letter, reminder post card, and other public use forms</li> <li>- Mailing of initial and reminder questionnaires</li> <li>- Mailing of advance notice letter, reminder post card, and other public use forms</li> </ul>	<p><b>\$419</b></p> <p>\$145</p> <p>\$27</p> <p>\$189</p> <p>\$58</p>	<p>For Census 2000, the Census Bureau has developed a new mailing treatment strategy (including an advance notice letter, first questionnaire, reminder/thank you postcard, and replacement questionnaire) that has been shown to increase mail response.</p> <p>In areas with a high concentration of people speaking a language other than English, the Bureau will provide questionnaires in the appropriate language.</p> <p>For Census 2000, the Census Bureau has conducted extensive research and testing to develop user-friendly questionnaires that make self-response and participation in the census easier.</p>

<p><b><u>Data Collection, Including Nonresponse Follow-up</u></b></p> <ul style="list-style-type: none"> <li>- Physical enumeration activities</li> <li>- ICM data collection</li> </ul>	<p><b>\$1,408</b></p> <p>\$1,114 \$294</p>	<p>The nonresponse follow-up operation in Census 2000 will incorporate the use of sampling to yield a 90 percent completion rate at the census tract level.</p> <p>In Census 2000, data collected in the physical enumeration and data collected in the ICM are integrated to produce a one-number census.</p> <p>Use of a service-based methodology to enumerate people with no usual residence in Census 2000.</p>
<p><b><u>Capturing Data from Returned Forms and Telephone Interviews</u></b></p> <ul style="list-style-type: none"> <li>- Data capture and processing oversight</li> <li>- Questionnaire receipt, check-in, editing, coding, and processing</li> <li>- Data capture center logistical activities</li> <li>- Processing system programming and support</li> <li>- Data capture system</li> <li>- Telephone questionnaire assistance and related activities</li> </ul>	<p><b>\$533</b></p> <p>\$29 \$99 \$48 \$232 \$76 \$49</p>	<p>Census 2000 data capture process will employ the latest commercially available electronic imaging technology i.e., taking an electronic photograph or "image" of each questionnaire and using the image to capture the data.</p> <p>The design, development, production, national deployment, and support of the entire data capture system has been contracted out to a single systems integrator.</p>
<p><b><u>Delivering Data Summaries to the Public</u></b></p> <ul style="list-style-type: none"> <li>- Prepare, review, and distribute census results</li> <li>- Prepare and distribute geographic products</li> </ul>	<p><b>\$166</b></p> <p>\$137 \$29</p>	<p>Development of an electronic data access and dissemination system and development/design of Census 2000 data products to meet user needs for faster and more timely access to census data. This replaces many of the electronic and paper products produced in 1990.</p>
<p><b><u>Evaluating the Results</u></b></p> <ul style="list-style-type: none"> <li>- Collect and evaluate information for use in next census</li> </ul>	<p><b>\$52</b></p>	<p>The Census Bureau will investigate potential methods for enhancing efficiency and effectiveness of future census activities.</p>
<p><b><u>Operations Management</u></b></p> <ul style="list-style-type: none"> <li>- Managing the census</li> </ul>	<p><b>\$86</b></p>	<p>Matrix management teams will generate efficiencies and lower staff levels at headquarters in Census 2000.</p>
<p><b>TOTAL</b></p>	<p><b>\$3,997</b></p>	

#### **IV. IMPROVEMENTS OF TRADITIONAL METHODS IN CENSUS 2000**

To ensure that Census 2000 will be both more accurate and more cost-effective than the 1990 Census, the Census Bureau has reviewed its procedures with input from a wide array of experts. The Bureau has asked these experts: Which parts of the process work best? Which can be done more effectively in some other way? Which can be eliminated? The result is an innovative departure from past practices that will substantially increase overall accuracy and address the differential undercount of children, renters, and minorities. At the same time, the new methods of enumeration will save money and deliver results more quickly. This section explains the improvements made to traditional census methods. Section V explains the improvements that involve scientific sampling.

##### **A. The Master Address File**

To conduct Census 2000, the Census Bureau must identify and locate an estimated 118 million housing units in the nation. The Bureau plans to accomplish this goal by developing and maintaining the Master Address File (MAF). This vital operation will take place with the assistance of the USPS, other federal agencies, tribal, state and local governments, community organizations, and by an intensive canvass of selected areas.

In 1990, whole housing units were missed often enough to contribute notably to the undercount problem. (See Section VIII(A)(1) for more detail.) Plans for Census 2000 are designed to address weaknesses found in the 1990 address list. In 1990, the Census Bureau relied on address lists purchased from vendors. It found that the purchased address lists were less accurate in low income areas because the lists were originally generated for marketing purposes. Vendors tended to focus their attention on wealthier, and therefore more profitable, areas. The 1990 experience also demonstrated the need to identify more carefully housing units in advance and place them on geographical maps.

The MAF being prepared for Census 2000 should be superior to the 1990 address list. The MAF will start with the USPS address list, a list that does not discriminate against certain areas because of their marketing potential. Partnerships with state and local officials, community organizations, and tribal governments will also play an important role in making sure the MAF is accurate as the local officials who know the areas best will help develop the MAF. Finally, the method used to create the MAF in rural areas will be superior because of intensive efforts well in advance of the census.

City-Style Addresses. The USPS uses the term “city-style” for an address such as “123 Main Street,” even though such an address occurs in small towns and increasingly along country roads. In areas where the USPS delivers mail primarily to city-style addresses, the Census Bureau will create the MAF by combining addresses from the 1990 Census Address Control File with those

addresses in the USPS Delivery Sequence File (DSF). The DSF is a national file of individual delivery point addresses. As part of a cooperative agreement, the USPS provides the Census Bureau with updated DSFs on a regular basis. The Bureau then locates these addresses in its computer mapping system called TIGER (Topologically Integrated Geographic Encoding and Referencing). If an address cannot be located, the location is researched and resolved through an office operation or through assistance from local partners. As a result of this research, the Bureau identifies new features and corrects and adds address ranges to the TIGER data base.

Non-city-Style Addresses. In late 1998 and early 1999, the Census Bureau will launch a comprehensive effort to canvass areas where most residences do not have city-style addresses. Over 30,000 canvassers will visit approximately 22 million residences without a street address to fix their locations on the TIGER system. The combination of innovative use of computer data and technology, along with these visits to areas without city-style addresses, will allow the Census Bureau to construct the most accurate address list ever, giving field enumerators more time to meet the other challenges presented by the 2000 count. The Bureau will conduct the initial data collection phase in these areas by having enumerators deliver addressed census questionnaires during an update/leave or an update/enumerate operation. Where there is no mailing address for the housing unit, or the mailing address is not a city-style address, the listing will include a location description.

The additional effort to identify and locate non-city-style addresses comes at a significant cost. City-style addresses are projected to cost \$1.40 per case compared to \$6.00 per case for non-city-style addresses.

Remote Areas. In a few extremely remote and sparsely settled areas, census enumerators will create the address list at the time of the initial census data collection while canvassing their assignment area and picking up or completing unaddressed questionnaires that the USPS previously had delivered to each household. The completed address listings and their geographic locations will be captured at this time.

Nontraditional Living Quarters. A separate operation will build an inventory of all facilities that are not traditional living quarters, for example, prisons and hospitals. The Bureau will interview an official at each location using a Facility Questionnaire. The responses to the questionnaire will identify each group quarters and any housing units associated with the location. The Bureau will classify each group quarters and any housing units at the location according to whether they will be enumerated as part of special place enumeration or through regular enumeration. The Bureau will add those group quarters and housing units to the MAF and link them to the TIGER data base.

Local Government Partnerships. The Bureau will rely on local knowledge to build the MAF. State, local, and tribal governments, regional and metropolitan planning agencies, and related nongovernmental organizations are encouraged to submit locally developed and maintained city-style address lists to the Census

Bureau to enhance the MAF. The participants will benefit by more complete and accurate data for their area. The Census Bureau will match the local list both to the MAF and TIGER data base.

The Census Bureau will attempt to verify the status of each newly identified address through ongoing matches to updated address information from the USPS, other independent sources, and its own field operations. Addresses that are not found on the TIGER system will be researched and resolved.

The LUCA (Local Update of Census Addresses) program is a partnership that will allow local and tribal governments to designate a liaison to review the portion of the MAF that covers their jurisdiction to help ensure its completeness. LUCA participants will benefit by more complete and accurate data for their area. Prior to the Census and after the initial targeting operations or address listing have been completed, the Census Bureau will send the liaison a listing from the MAF and the accompanying maps for their jurisdiction. The liaison will review the addresses and provide the Census Bureau with updates (adds, deletes, and corrections). After processing the LUCA input, the Census Bureau will provide feedback on the status of the adds, deletes, and corrections to the liaison. The updated address list then will be used to deliver census questionnaires (either by mail or by an enumerator's visit).

## **B. New Outreach Methods**

Several innovations are planned for outreach. Collectively, these new methods should increase response rates beyond those expected with the 1990 methodology.

Multiple contacts. For the first time ever, Census 2000 will implement a multiple mail contact strategy. Instead of mailing just one questionnaire, the Census Bureau will mail two waves of questionnaires, each preceded by a mailed notice/reminder. This strategy has paid big dividends for the private sector and has proved effective in Census Bureau tests. The Census Bureau projects that a single mailing would result in further erosion of the response rate to 55 percent, but that its multiple notices and questionnaires will boost the response rate to 67 percent.

More ways to respond. In 1990, respondents had to find their form in the mail; in 2000, the forms will find respondents. The Bureau will make forms available in stores and malls, in civic or community centers, in schools, and other locations frequented by the public. A well-publicized toll-free telephone number will be available for those who wish to respond to the census by phone. In remote or sparsely-populated areas, enumerators will visit each housing unit and pick up or complete unaddressed short form questionnaires and administer the long form at predesignated sample households.

Multiple languages. In 2000, as in all prior decennial censuses, questionnaires will be in English. For the first time in a decennial census, however, some households in Census 2000 will receive two questionnaires -- one in English and one in another language. Specific neighborhoods known to have a high proportion of households

more familiar with languages other than English will be sent questionnaires in their second language as well as a questionnaire in English. Forms in other languages will also be made available in locations frequented by non-English speakers. The Census Bureau has made Spanish-language questionnaires available in the past, but questionnaires in languages other than English have never before been included in the initial mail-out package.

### **C. New Technology**

The Census Bureau plans to introduce several new technologies in Census 2000:

Unduplication. Modern technology allows the Bureau to spot and eliminate multiple responses from the same household. One of the main goals of Census 2000 is to make it simpler for people to be counted by having census forms available in public locations, provided in multiple language translations, and mailed out twice in mail-out/mail-back areas. Responses to the census also will be accepted over the telephone and possibly on the Internet. Providing these response options will make it easier for everyone to be counted, but will increase the possibility of multiple responses for a given person and household. A complete, accurate address list, high speed data capture capabilities, along with automated matching technologies will be the keys to avoiding the duplication of people and residences. Unduplication of multiple responses was not feasible or necessary in past Censuses because available technology and costs permitted only a small subset of person names to be data captured and unduplicated.

Data Retrieval Technology. Through contracting with private vendors, the Census Bureau will utilize the best available data capture methodology in Census 2000. The Bureau has successfully tested the hardware and software that converts handwriting on the questionnaire into computer form with minor editing by a technician. The plan for data capture will use off-the-shelf hardware and software to record a full electronic image of every questionnaire. It will sort mail-return questionnaires automatically to ensure timely conversion and capture of critical information needed prior to nonresponse follow-up. It will use Optical Mark Recognition (OMR) to capture all check-box items, Intelligent Character Recognition (ICR) and key-from-image to capture write-in character-based data items. The Bureau will also conduct quality assurance on all data capture activities. The system will reduce the logistical burdens associated with handling large volumes of paper questionnaires. Once original questionnaires are checked-in, prepared and scanned, all subsequent operations will be accomplished using the electronic image and captured data.

## **V. USE OF SCIENTIFIC SAMPLING TO INCREASE ACCURACY**

In our common experience, “sampling” occurs whenever the information on a portion of a population is used to infer information on the population as a whole. We use samples every day to characterize a larger group -- for manufacturing quality checks, for medical tests, for determining air and water quality, and for conducting audits, to name a few. In laymen’s terms, a “sample” is taken whenever the whole is represented by less than the whole. Among professional statisticians, the term “sample” is reserved for instances when the selection of the smaller population is based on the methodology of their science. The sampling proposed for Census 2000 is scientifically based; improves accuracy; eliminates the traditional undercount of children, renters and minorities; and saves money.

### **A. Reliance on Sampling in Previous Censuses**

In the debate over methods to be used in Census 2000, **the issue is not whether to “sample” but whether to sample scientifically.** Census takers have never been able to contact and count each and every resident of this nation. As a result, information on less than the whole population has always been used to characterize the whole population.

Census 2000 will not be the first time that the Census Bureau has used statistical methods to correct for problems in physical enumeration and to provide a more accurate final result. Since at least 1940, statistical imputation has been used when an enumerator knew that a housing unit was occupied, but could not obtain information on the number of people living in that unit. In 1980, statistical imputation raised the physical enumeration total by 761,000 people. The number and rate of people imputed in the 1990 Census was only 53,590. Automated data control systems and field procedures may have discouraged enumerators from turning in incomplete questionnaires. In 1970, the Census Bureau used sampling to impute people to addresses that had initially been assumed vacant. The sample of 13,546 housing units initially presumed “vacant” found that 11.4 percent of them should be reclassified as “occupied.” The National Vacancy Check added 1,068,882 people, or 0.5 percent of the total, to the 1970 Census.

Apart from the population totals, the Bureau has historically used statistical methods extensively to make up for incomplete census information. For example, information is asked about each individual’s age, sex and race. Established statistical methods were used to infer missing information.

In other efforts, the Bureau has used statistical methods to represent the whole population when less than complete responses are obtained. For example, the Census of Industries every five years makes an effort to contact every large company in specific industries. Some companies do not respond and statistical methods are used to account for them in the totals ultimately published.

## **B. Support for Sampling within the Scientific Community**

The assumption underlying the traditional census method has been that the most accurate representation of the entire population would come from an intensive effort to physically contact every individual household. The experience of 1990 proved that this assumption was no longer valid and that to “pour more money into traditional methods” was not satisfactory. Because of changes in our society, a sample drawn by including only those physically contacted became markedly inaccurate. After 1990, a scientific consensus emerged that, while we should continue to pursue physical contact with every household to supply information, we should use the best statistical science to organize collection of information on those who fail to respond by mail or phone. In other words, the old system is no longer adequate in light of societal changes.

Census 2000 will use sampling in two new ways: (1) to follow up on housing units that do not respond by mail or phone; and (2) to use ICM to minimize the coverage error associated with past censuses. In addition, Census 2000 will sample housing units considered vacant by the USPS.

These methods have been endorsed by three Academy panels and by numerous other organizations:

- The American Statistical Association, a 157 year old group with more than 19,000 members nationwide, convened a “Blue Ribbon Panel” to assess the use of sampling in Census 2000. In its 1996 report, that panel “endorsed the use of sampling,” concluding that it is “consistent with best statistical practice.” It notes that “sampling is used widely in science, medicine, government, agriculture, and business because it is the fundamental basis for addressing specific questions in these areas. Sampling is a critical tool for reducing uncertainty.” The report observed that “[s]pecific areas that use statistical sampling extensively include auditing, market research, quality assurance, approving new drugs, and medical testing ... Sampling permits observations to be made efficiently, economically, and fairly.”
- The American Sociological Association, founded in 1905, is a national professional society of 12,500 sociologists, research scientists, and others interested in research, teaching, and application of sociology. On January 25, 1997, the American Sociological Association unanimously approved a resolution supporting the use of sampling in the decennial Census. This resolution strongly urged “the Secretary of Commerce and Congress to support unequivocally the use of sampling for non-responding households and for reducing the differential undercount in the 2000 Census.”
- The General Accounting Office also favors the use of sampling, stating that it is “encouraged that the Bureau has decided to sample those households failing to respond to Census questionnaires rather than conducting a 100-percent follow-up as it has in the past.” “Sampling households that fail to respond to questionnaires produces substantial cost savings and should improve final quality.”



- The Inspector General of the Department of Commerce has repeatedly endorsed the plan to use sampling in Census 2000. Most recently, in a May 5, 1997 letter to Senator Stevens, Chairman of the Senate Committee on Appropriations, the Inspector General stated that, “[o]ver the past two years, we have issued reports, testified, and briefed bureau, departmental, and congressional principals and their staff members on our support for the use of statistical sampling in the 2000 Census. We continue to believe that, if carefully planned and implemented, sampling can be employed by the bureau in the 2000 Census to produce overall more accurate results than were produced in the 1990 Census, at an acceptable cost.”

### **C. Sampling to Collect Long Form Data**

The Census Bureau has used sampling techniques since 1940 to collect some of the most important decennial census data. Prior to the institution of the short form the Census Bureau had asked detailed census questions of every resident. In fact, the Census Act mandates the use of sampling in the decennial census, *see*, 13 U.S.C. §§ 141, 195. As it has in the last six decennial censuses, the Census Bureau will deliver the long form questionnaire to a sample of housing units.

In 2000, the long form will ask the same 7 questions that appear on the short form, plus questions on an additional 27 subjects that are either specifically required by law to be included in the census or are required to implement other federal programs and the decennial census is the only source of the data. Using sampling to collect long form information will enable the Census Bureau to meet the objectives of controlling cost and maintaining or reducing respondent burden.

The long form is a cost-effective tool for gathering information to evaluate and implement federal programs. Dozens of agencies depend on the long form for the information they need to run their programs, including the Department of Defense, the Federal Reserve, the National Center for Health Statistics, the Department of Labor and many more. The following are just a few examples of how long form data is used:

- Federal and local emergency management agencies use census data to assess the amount of displacement caused by earthquakes, hurricanes, floods and other natural disasters.
- Planners must have information about where people work, where they live, how they get to work, and when they leave for work to build roads, tunnels, and bridges in areas that need them.
- The Department of Veterans Affairs uses census data on age, veteran status, period of service, years of service, and residence five years ago to determine where hospitals, nursing homes, and other services should be located.

The Census Bureau will use a variable rate sampling scheme in Census 2000 to collect long form data. The variable rate sampling scheme for Census 2000 will probably be similar to the 1990 sampling scheme:

- The overall sampling rate will be about one in six, or 17 percent.
- The sampling rate in general purpose governmental units with populations fewer than 2,500 will be one in two. Designated American Indian and Alaska Native areas will also receive a one in two sample.
- The sampling rate in other governmental units will be one in six or one in eight.

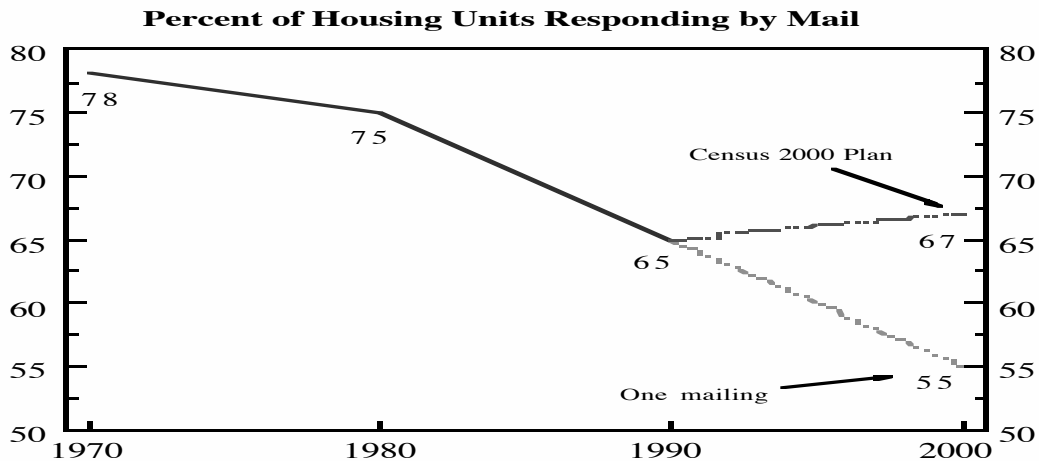
Variable rate sampling will allow the Census Bureau to allocate the sample efficiently while reducing respondent burden and maintaining the accuracy and reliability of census data at small geographic levels (census tracts, and small communities).

#### **D. Sampling in the Postal Vacancy Check**

As noted in Section V(A), the Census Bureau used sampling in 1970 in response to concerns that too many housing units had been erroneously identified as vacant. The Census Bureau estimates that the USPS will identify about five percent of all housing units as vacant in 2000. In order to correct for anticipated errors in this designation, the Census Bureau will send interviewers to one out of every ten of the housing units that the USPS indicates are vacant. The number of housing units that are found to be occupied and the number of people living there will be used to estimate the total population of units initially designated as vacant. Scientifically proven statistical methods will be used to account for the total population and the estimated number of occupied and vacant housing units in the postal vacant universe.

#### **E. Sampling for Nonresponse Follow-up**

In conducting Census 2000, the Census Bureau will rely mainly on mail returns of census questionnaires, as it has in every census since 1970. If every housing unit returned its form by mail, the Bureau would not need a nonresponse follow-up operation, but the mail response rate declined markedly from 78 percent in 1970 to 65 percent in 1990. The Census Bureau estimates that mail response in 2000 would fall to 55 percent with one mailing, but can be raised to 67 percent with its current plan for two waves of notices and questionnaires and other innovations. That will leave 34 million occupied housing units not expected to respond.



The Bureau will attempt to contact these nonresponding housing units in the nonresponse follow-up portion of Census 2000. The Bureau's plan for this operation includes a limited use of statistical sampling to assure that data are collected from at least 90 percent of the housing units in each census tract. (A census tract is a neighborhood with roughly 1,700 housing units and 4,000 people. Tracts are designed to have homogeneous population characteristics, economic status, and living conditions. There will be more than 60,000 census tracts in 2000.) All questionnaires mailed back during the data collection period will be included in the enumeration process.

The most difficult logistical segment of Census 2000 becomes more manageable with a scientific sample of nonresponding housing units. Enumerators must visit 22.5 million housing units rather than the 34 million housing units they would need to visit without sampling. Reducing the number of housing units to be visited will allow the Census Bureau to hire fewer and better qualified enumerators. And the time and effort that would have been spent recruiting, screening, training and managing additional temporary employees can be spent on meeting the other challenges involved in Census 2000.

By reducing the burden on enumerators, sampling for nonresponse follow-up will help ensure that the ICM can begin on time. The longer the delay between Census Day and the ICM, the more respondents are likely to provide inconsistent responses (out of forgetfulness, or because of the continuous turnover in housing units -- which affects approximately 150,000 housing units each month).

In its second Interim Report on June 10, 1997, the Academy's Panel on Alternative Methodologies commented on the benefits of sampling:

... [W]e do not believe that a Census of acceptable accuracy and cost is possible without the use of sampling procedures, both for non-response follow-up and integrated coverage measurement ... [T]he use of sampling will reduce the field workload and may result in more timely completion of the nonresponse follow-up procedures in the field.

Selecting the Sample for Nonresponse Follow-up. The addresses that will be part of the nonresponse sample will be evenly distributed across all addresses in each census tract not returning forms by mail or answer by telephone. The Census Bureau will achieve this goal by using scientifically-proven techniques, including computer-generated random sampling, to make sure that every nonresponding address in each census tract has an equal chance of selection. Enumerators will be given a list of specific addresses of nonresponding housing units within a census tract as soon as the mail-in phase is complete. Because these addresses are part of a random sample, they will be statistically representative of all housing units in that nonresponding tract.

To obtain information from 90 percent of housing units in each census tract, those tracts with lower mail response rates will have a higher share of housing units sampled. For example, for census tracts in which 80 percent of addresses return their forms by mail, the enumerators will be assigned randomly-selected addresses that represent half of the addresses that did not respond. If only 70 percent of addresses return their forms, enumerators will be assigned two of every three addresses not responding. And if the response rate is lower, enumerators will contact an even greater share of nonresponding addresses -- more than adequately covering all segments of the census tracts not responding.

Response Rate	Size of Sample	Total Direct Contacts
30 percent	6 in 7	90 percent
40 percent	5 in 6	90 percent
50 percent	4 in 5	90 percent
60 percent	3 in 4	90 percent
70 percent	2 in 3	90 percent
80 percent	1 in 2	90 percent
90 percent or more	1 in 10	91 percent

Note: The Census Bureau is reviewing the June 1997 recommendation of the Academy panel to increase the sample for census tracts with high response rates.

Interviewing by Temporary Field Staff. The nonresponse follow-up operation is the largest single operation in Census 2000. In order to follow up with housing units not returning questionnaires in the mail, the Census Bureau will recruit, hire, train and supervise a massive temporary field staff. Since virtually every urban and rural community and neighborhood has housing units requiring follow-up, the Census Bureau must mount a nationwide recruiting campaign. Temporary staff are hired for this operation because it must be completed in only a few weeks. The most productive times to make contact with households are evenings and weekends, so this staff does not work a full-time 40-hour work week. Most importantly, the Census Bureau tries to assign this temporary staff to

neighborhoods with which they are familiar so that they are better received and the likelihood of collecting accurate data is enhanced.

Quality Assurance for Nonresponse Follow-up. The Census Bureau implements quality assurance operations for each of its major operations to insure that the results meet acceptable standards of quality. The quality assurance operation for nonresponse follow-up includes a re-interview of a portion of the cases completed by nonresponse follow-up enumerators. In this re-interviewing, an independent staff member conducts a brief interview to insure that the household was directly contacted the first time.

The Estimation Procedure. The characteristics of the sample housing units will be used to estimate the characteristics of the housing units not in the nonresponse follow-up sample. Take, for example, a census tract with 1,000 housing units and mail-back responses from 800 (80 percent). In that case, information on the remaining 200 housing units would be based on a one in two sample of 100 housing units. If that same tract had responses from only 400 (40 percent), the Census Bureau would interview 500 addresses to estimate the 600 nonresponding addresses.

## **F. Integrated Coverage Measurement**

Of all the innovations to improve accuracy in Census 2000, the most critical is Integrated Coverage Measurement (ICM). Inaccuracy largely stems from two problems. First, some housing units are never contacted because they are missing from the address list. The Census Bureau's considerable effort to improve the quality of the address list for both urban and rural areas should serve to reduce the number of missed housing units. The second and much larger source of inaccuracy comes from missing people in housing units that do supply some information. The ICM not only helps with holes in the address list, it represents an effective way to address the second problem. That is why all three National Academy of Sciences panels recommended inclusion of the ICM in the plan for Census 2000.

The Census Bureau has a great deal of experience with Dual System Estimation, the methodology to be used in the ICM. The methodology was employed in the past two censuses to evaluate census quality. The methodology has undergone substantial review and improvement by the Census Bureau, the National Academy of Sciences, and by experts in statistical methodology from across the country. ICM methodology is generally accepted as the most reliable method to improve census results.

To conduct the ICM, Census Bureau enumerators will interview a carefully-selected random sample of about 750,000 housing units. This sample will be selected to include blocks from all areas of the country, with all race and ethnic groups, from all sizes of towns and cities, and from rural areas. The objective is to determine what proportion of the people living in the sample blocks were included and what proportion were excluded in the initial phases of the census.

Because this sample is very large, and drawn separately for each state, it will provide reliable population numbers for every state and Congressional district.

Selecting the ICM Sample. As was done for the 1990 PES, to select the ICM sample for 2000, the Census Bureau plans to classify each of the country's seven million blocks into groups known as strata. These strata will be based on the characteristics of each block in the 1990 Census, such as the block's state, racial and ethnic composition, and proportion of homeowners to renters. The Census Bureau expects the types of sampling strata to be similar to those used for the 1990 Post Enumeration Survey (PES).

In Census 2000, the sample will be much larger (750,000 housing units versus about 150,000 in 1990) and strata will be defined within each state (strata crossed state lines in 1990). This will mean that in 2000, sample data from one state will not be used to determine the population total in a different state. In 2000, every state will have its own ICM sample, representative of all parts of that state. An example of a homogenous sampling stratum would be: All blocks in large central cities with a 1990 Census population that was 30 percent or more African American renters and with 10 percent or more Hispanic renters.

The Bureau will then select blocks at random from each stratum, for a total of 25,000 blocks. With blocks having an average of 30 housing units, the ICM will obtain information from 750,000 housing units. This process will establish a sample that is large enough, and sufficiently representative, to estimate population totals for each state. By stratifying and weighting the sample blocks the ICM will address coverage errors for specific population groups or areas, even if the individuals in the ICM sample constitute a relatively small part of the total population of the nation.

The ICM Address List. In order to ensure accuracy, the address list for the ICM is initially developed without use of the Master Address File. The Census Bureau will conduct a thorough, labor intensive canvass of each block in the ICM, an effort that would be logistically impracticable and too costly to repeat for all blocks in the entire nation. The list of addresses from the canvass effort is then matched with the MAF and differences are resolved.

The Initial ICM Interviews. Enumerators will use the ICM address list to conduct interviews at the 750,000 housing units in the sample blocks, and thereby establish an independent roster of Census Day residents. The enumerators will administer the ICM questionnaire and enter data via laptop computers. The Bureau expects to hire about 12,500 ICM interviewers and probably another 4,500 supervisors and quality assurance representatives.

Follow-up ICM Interviews. All housing units in which discrepancies are detected between the pre-ICM response and ICM response are designated for reconciliation and are usually assigned to a follow-up interview. The follow-up interviewer revisits each address for which there are inconsistent results and attempts to find the correct answer or the "true" situation. This process leads to a

determination of whether the ICM response or the initial phase of the census is correct for a particular unit.

Poststrata. Each person is then assigned to a unique poststratum, or group of people who have similar chances (probability) of being counted in the initial data collection operation. The poststrata are defined by state geographic subdivision (such as rural or urban), owner or renter, age, sex, race and Hispanic origin. The 1990 Census Post Enumeration Survey used 357 poststrata to characterize the population of four geographic regions of the United States. The ICM process in Census 2000 will also use poststrata to produce estimates for each state. The results of the 1990 PES found that people living in rental housing units were much more likely to be uncounted. Therefore, the revised undercount estimates published in 1992 used owner-renter status in all the poststrata, except for Asians and Pacific Islanders, and American Indians on reservations. Basing the poststratification on the combination of variables defined by owner-renter status and race-Hispanic ethnicity improved the population estimates.

Nonresponse in the ICM. If the Census Bureau is unable to obtain an interview with an ICM sampled housing unit, despite intense effort, the Bureau will use imputation to account for the household. Imputation is a method that the Census Bureau has traditionally used when faced with legal deadlines and no alternatives to account for that household. Census Bureau research (from the 1995 Census Test and from the 1990 Census) indicates that imputation will have no major effects on the final results and that imputation is more accurate than leaving out missing information.

Dual System Estimation. Comparing the results of the ICM with the results of the initial phase of the Census will reveal who was missed in the sample blocks. The Bureau will then use a statistical method called Dual System Estimation to estimate the extent to which housing units and people were correctly included in the initial data collection phase, missed, or counted in error for each state. Dual System Estimation is a widely-known and accepted statistical method that provides an accurate accounting of populations. Dual System Estimation works by comparing two independent sets (dual systems) of information on the same geographic area. In this case, the population being measured is in the 25,000 blocks that compose the ICM sample and the goal is to determine the true population of these blocks. For ICM sample blocks in Census 2000, the non-ICM set of information will consist of all direct responses gathered by mail, telephone, and personal visits. The second set of information comes from the ICM data on those sample blocks. Taken together, the two systems produce a single estimate of the total population.

Most of the housing units in the sample blocks and the people residing in them will be in both the initial phase and the ICM. A few people will be in one measure but not the other. Occasionally, an entire housing unit will be in one measure but not the other. Comparing the results of the ICM to the results from the initial effort (including mail, telephone and personal interviews) will inform the Census Bureau of the proportion of the population included in both, the proportion missed in the non-ICM effort, and the share included in the non-ICM effort but not in the

ICM. These results are used to produce an estimation factor for each poststratum. Estimation factors account for the differences between the two efforts and for cases missed in both systems. The estimation factors are applied to the initial phase to estimate the total population and housing units in each poststratum. The sum across poststrata is used to estimate state totals. State totals are summed to national totals. The state level estimation factors by poststrata are used to produce population and housing unit estimates for every block in the nation.

The Census Bureau is confident in the Dual System Estimation methodology based on its experience implementing Dual System Estimation and its expertise analyzing and explaining Dual System Estimation results. In 1990, the Census Bureau used Dual System Estimation to produce coverage estimates; the results from the 1990 Census were consistent with the independent benchmark of Demographic Analysis.

The Dual System Estimation theory requires that the two systems collect data independently. It does not require the ICM be superior to the “initial” system. In practice, the size of the sample will allow the Census Bureau to collect data in ways that would be prohibitive if attempted on the entire population. The ICM provides an independent, intensively-researched address list, and an intensive personal interview, designed to elicit complete and accurate information on people with characteristics that typically are missed in enumeration.



## **VI. OPTIONS FOR ADDRESSING AREAS WITH HIGH UNDERCOUNT RATES IN THE ABSENCE OF SCIENTIFIC SAMPLING**

The Census Bureau has designed its plan to achieve the most accurate enumeration possible within the constraints of established statistical procedures, available and reliable technology, reasonable budget resources, and the conditions of modern American society. Some have argued that the Census Bureau should be prohibited from using established and widely-recommended statistical procedures, including sampling, and should rely solely on physical enumeration methods. To rely entirely on physical enumeration would result in a substantially less accurate census even though it would cost substantially more.

The Census Bureau believes that, without the introduction of a limited use of sampling, Census 2000 will be even less accurate than the 1990 Census. The population expected to have a high rate of undercount with traditional methods has grown more rapidly than the total population. As a result, the Census Bureau estimates that a physical enumeration in 2000 would fall short of the actual population by at least 1.9 percent -- more than 5 million people. Moreover, conducting Census 2000 without sampling would not address the recurring problem of the differential undercount.

### **A. The Census Bureau Could Face Severe Labor Difficulties if Forced to Take Census 2000 without Sampling**

To require that Census 2000 be conducted without sampling could raise practical hiring difficulties as well as increasing cost. Just as changing conditions have reduced mail response rates, so have changing conditions reduced the Census Bureau's ability to attract and hire qualified enumerators. Using sampling techniques to complete the nonresponse follow-up operation significantly reduces the number of temporary enumerators that must be hired. Attracting and hiring a sufficient number of temporary employees in 2000 will be a difficult undertaking, even with the use of sampling.

The plan for Census 2000 assumes that the Census Bureau would have to hire over a quarter of a million temporary employees within a very few months. To do so, the Bureau would have to recruit over three million individuals. The nonresponse follow-up operation is particularly labor intensive. At peak employment, about 117,000 workers, approximately 47 percent of all temporary employees for Census 2000, would be hired to conduct the nonresponse follow-up at 22.5 million homes. Without sampling, enumerators would have to visit an additional 12 million homes, requiring 59,000 additional enumerators for the nonresponse follow-up operation. The Census Bureau has retained the services of Westat, Inc. to help it develop a model for setting enumerator wage rates in 2000. Wages will have to be set high enough to attract and *retain* qualified employees.

**B. The Only Alternative to Sampling Would Be to Rely More Heavily on Traditional Methods, Methods with Proven Limits**

The Census Bureau agrees with the Academy's conclusion that more radical alternatives to a traditional enumeration (a national register, an administrative records Census and a Census conducted by the U.S. Postal Service) are either not feasible or not consistent with American values.

As discussed in Section III, the Census Bureau is committed to using traditional enumeration methods and the plan for Census 2000 contains several innovations to these traditional methods. The Census Bureau's experience, and the Academy's research, however, indicate that even with improvements traditional methods of enumeration cannot achieve satisfactory accuracy.

Between 1970 and 1990, the Census Bureau tested a number of outreach, coverage, and collection procedures designed to increase accuracy, particularly in areas with high undercount rates. Some of these innovations, described below, improved accuracy -- though they did not prevent the level of inaccuracy from rising in 1990 -- and will be used in Census 2000. Other innovations were proved to be prohibitively expensive, hard to control, error prone, or ineffective.

Among the innovations that have been tested are:

*Advertising* -- The Census Bureau launched a massive outreach campaign in the 1990 Census, using public service announcements as advertisement. The Bureau's conclusion was that advertising improved the mail response rate but that its effectiveness in reaching areas with high undercount rates needed improvement.

Status: In Census 2000 the Bureau plans a \$100 million paid advertising campaign targeted at areas with low mail response rates. (See the discussion in Section II(B)(2).)

*Advance Post Office Check* -- USPS letter carriers in 1990 verified the completeness of the address list by making corrections, identifying duplicate and undeliverable addresses, and reporting missing addresses. This approach was successful, but not efficient enough.

Status: Better partnership with the USPS and with state, local, and tribal governments have allowed the Bureau to replace this procedure with more efficient and more comprehensive programs for Census 2000. The 1990 Census spent too much time and money developing an address list that the USPS already had assembled. For Census 2000, the Census Bureau began with the USPS information to avoid a costly duplication of effort. State, local, and tribal governments are working with the Bureau to correct and update census maps on a continuing basis.

*Casing Check* -- Immediately prior to the delivery of the 1990 Census questionnaires, USPS letter carriers identified deliverable and undeliverable

addresses, and notified the Bureau of any homes on their route that did not appear on the Bureau's list.

Status: Better partnership with the USPS, and with state, local, and tribal governments have allowed the Bureau to replace this procedure with more efficient and more comprehensive programs for Census 2000.

*Census Awareness and Products Program* -- This program built awareness about the 1990 Census by educating the public and encouraging it to participate. Census Community Awareness Specialists were hired to contact teachers, mayors, religious groups, and others, and to develop promotional materials.

Status: This program forms the basis of the Census 2000 partnership programs.

*Census Closeout Address Check* -- During the final stages of field follow-up activities in 1990 USPS letter carriers provided information about the type of structure, occupancy status on Census Day, and the number of Census Day occupants for unenumerated units. This approach was successful, but not efficient enough.

Status: Better partnership with the USPS, and with state, local, and tribal governments have allowed the Bureau to replace this procedure with more efficient and more comprehensive programs for Census 2000.

*Parolee/Probationer Coverage Improvement Program* -- People on parole or probation completed a unique census form to help ensure that they were counted in the 1990 Census.

Status: The Census Bureau dropped this program from consideration for Census 2000 because it was error prone and not cost effective.

*Precensus and Postcensus Local Review* -- In 1990 local and tribal government officials reviewed housing unit counts and group quarters population counts for each census block in their jurisdictions and identified missed units.

Status: This program has been replaced in Census 2000 by the greatly expanded Local Update of Census Addresses (LUCA) program, made possible by amendments to Title 13.

*Recanvass* -- Enumerators did a second canvass of addresses in selected neighborhoods to look for missed units in areas with evidence of deficient housing unit counts.

Status: The Census Bureau dropped this program from consideration for Census 2000 because it was ineffective.

*Shelter and Street Enumeration* -- The Census Bureau took a special enumeration of people in shelters and at pre-identified street locations in the 1990 Census.

Status: In Census 2000, the service-based enumeration operation will use statistical estimation to improve the enumeration of people with no usual residence. ICM procedures will be adapted to conduct an initial enumeration at locations where people with no usual residence receive services (shelters and soup kitchens). A second visit may be conducted at a sample of locations to account for people who were not present at the time of the initial data collection operation, but who do use services at other times. If sampling and estimation were prohibited, the Census Bureau could not make this follow-up visit. People who were not present at the time of the initial data collection operation would be encouraged to use other opportunities to be included in the census but coverage of the people who frequent these sites would not be as complete.

*Telephone Assistance Adds* -- In the 1990 Census, people who called to say they had not received a questionnaire in the mail were told to wait for a personal visit interview.

Status: The Census Bureau will expand the telephone assistance program in Census 2000 to include the ability to take interviews over the phone.

*Transient Enumeration* -- The Census Bureau took a special enumeration of individuals residing in hotels, motels, tourist homes, campgrounds, and marinas in the 1990 Census.

Status: This program has been retained for Census 2000.

*Vacant/Delete/Movers Check* -- Enumerators revisited all addresses classified during nonresponse follow-up as vacant or delete to verify Census Day occupancy status and to complete questionnaires for people who moved during the enumeration period.

Status: This operation will be done more efficiently in Census 2000 by revisiting only a sample of the units classified as vacant. Without sampling, that efficiency would be lost.

*Were You Counted Campaign* -- People who thought they were missed in the 1990 Census had the opportunity to complete a simplified census questionnaire late in the census program.

Status: In 2000, the "Be Counted" program will make census forms available in public places, such as community centers and post offices -- at the same time as key data collection activities and advertising are taking place.

*Post Enumeration Post Office Check* -- An operation to improve coverage in very rural areas was used in the 1970 and 1980 censuses. After the census enumeration was completed, the Postal Service reviewed the addresses collected by the enumerators and identified any missed living quarters.

Status: This approach was not successful enough to merit its continuation.

**C. Spending More on Outreach Instead of Sampling Would Leave an Unacceptably Large Undercount and Have Biased Results**

Spending more on public outreach and enumerators will not adequately address the undercount problem. For a variety of reasons, accounting for every address does not guarantee that every person at that address is reported. In some cases, nontraditional family situations make reporting difficult. For example, children who share time with divorced parents, or with their extended family, may not be counted in either household. In other cases, non-traditional housing situations may lead to incomplete results. For example, landlords may assume that a tenant sharing their home is being counted separately. In some cases, individuals are missed because the respondent could not list his or her entire family in the space provided, or because the respondent's English language skills are limited. In others, individuals are missed because of enumerator error.

**D. Without Sampling, Costs Would Increase by at Least \$675 Million and the Final Count Would Be Less Accurate than the 1990 Census**

As planned, Census 2000 is projected to cost approximately \$4 billion. A cornerstone of the Census Bureau's plan is to reduce costs and increase accuracy by using scientific sampling methods while making a "best faith effort" to include every resident of the United States with traditional enumeration methods. The use of statistical methods, both to complete nonresponse follow-up and to implement the ICM survey, will address historic problems associated with cost, incomplete coverage, and the differential undercount. Using sampling techniques to finish the initial phase will provide the time and funds needed to implement the ICM survey, which will eliminate the need for costly add-on operations that otherwise would be necessary to increase the completeness of Census 2000. If sampling is prohibited while these add-on activities are implemented, Census 2000 would probably leave approximately 1.9 percent of the population uncounted and thus be even less accurate than the 1990 Census.

The following sections analyze available options for reducing the undercount in areas with high undercounts, without the use of statistical methodologies. The increased costs, which are summarized at the end of this Section, would be attributable to the following additional expenditures:

*100 Percent Follow-up On Nonresponding Units.* After the mail and telephone response period, the plan for Census 2000 calls for collecting information from a sample of nonresponding addresses in each Census tract. The sample will be drawn to assure that information will be obtained from at least 90 percent of addresses in each census tract. This effort will require sending enumerators to 22.5 million addresses. However, with a ban on sampling, enumerators would have to visit all of the expected 34.5 million nonresponding addresses. To send enumerators to an additional 12 million nonresponding addresses would cost a projected \$400 million more.

*Significantly More Effort Would Be Required to Verify Vacant Units.* Since 1970, letter carriers have been the primary source of information as to which housing

units are vacant. Because the 1970 Census found 11.4 percent of its sample of vacant units were in fact occupied, the Bureau had to raise the enumeration level by more than one million persons. In the 1990 Census, the Census Bureau initially classified about 7 million housing units as “vacant” and 3 million addresses as “not living quarters” and assigned them for field verification. About 9 percent of the initially “vacant” units had to be reclassified as “occupied” on Census Day and about 12 percent of the units classified as “not living quarters” changed classification after field verification. In total, this effort resulted in counting 1.5 million more people.

The plan for Census 2000 includes a field verification of a ten percent sample of the cases identified as vacant by the Postal Service. The Bureau expects that the USPS will identify five percent of housing units as being vacant on April 1, 2000. This verification will ensure the integrity of the information provided by the USPS, and gather information about the characteristics of those vacant units. If sampling were not permitted in Census 2000, the additional amount of effort and person hours needed to verify occupancy status would dwarf the considerable effort now planned. Under a 100 percent verification plan, the Census Bureau would assign all of these cases to field staff to verify their status as of Census Day. The field visits will be integrated with nonresponse follow-up, starting in April 2000. This additional effort would require 25,000 to 30,000 enumerators, as opposed to the 5,000 required using sampling, at an additional cost of \$200 million.

*100 Percent Follow-up of Incomplete Questionnaires Would Be Necessary.* The plan for Census 2000 requires only a computer check of the questionnaires that are returned by mail for evidence that coverage problems exist. Questionnaires flagged as having problems are sent for a telephone follow-up to attempt to resolve the discrepancies. Estimates are that approximately 0.5 percent of questionnaires will require follow-up. The ICM will account for coverage errors not corrected by this operation. The plan also includes follow-up for households containing more than five people (since the questionnaire only provides space for recording data for five people).

A plan that does not include sampling would require a much more thorough editing and follow-up of the questionnaires. During the processing of mail returns, phone-in responses, and check-in of enumerator-completed questionnaires at Local Census Offices, the Census Bureau would identify all cases that require additional contact to ensure the accuracy of the reported information. These cases would include households that return a blank questionnaire in the mail, questionnaires with a discrepancy between the total number of household members and the number of people for whom census data are provided (for example, the questionnaire lists the names of four persons but provides information only for two), and questionnaires with other indications of coverage problems, such as confusion over residency status. This additional effort would cost \$150 million more than current plans.

*Expanded Partnership and Promotion Activities Would be Required.* The promotion and outreach program is designed to motivate people to respond. While the Census Bureau does not envision additional partnership and promotion activities if it is not allowed to use sampling, the time period for planned promotion activities would have to be extended to include a longer nonresponse

follow-up period. Also, the Census Bureau would need to hire additional staff to provide at least one partnership specialist for each Local Census Office. The Bureau would expect to pay \$75 to \$150 million more for these activities.

*Deployment of Special Enumeration Activities.* The Census Bureau has developed targeted methods to supplement its basic data collection strategy. If the Bureau is banned from using the best available statistical methods in 2000, the activities summarized below would need to be intensified, and started sooner than currently planned. They would cost \$25 to \$50 million more.

Team enumeration. In targeted areas, a team or crew of enumerators conducts the enumeration in a short period of time. Team enumeration will be used in areas where conditions in the field may interfere with the timely completion of the enumeration. These conditions may be high concentrations of multi-unit buildings, enumerator safety concerns, and low enumerator production rates.

Urban update/leave methodologies. The urban update/leave operation will be conducted in selected urban areas where mail delivery is a problem. During the operation, enumerators in teams will hand-deliver census questionnaires to households and ask respondents to complete the forms and mail them back.

*Quality Assurance.* The techniques used for quality assurance operations would need to be enhanced. Without ICM to provide final quality assurance, all quality assurance operations will require extra efforts at an extra cost of \$25 to \$50 million.

*Postcensus Evaluation Study.* The Census Bureau will conduct a thorough evaluation of Census 2000. Without the use of sampling the Bureau would gather data for the evaluation through a post enumeration survey.

#### **Summary of Additional Costs with a Ban on Statistical Methods**

100-Percent Follow-up on Nonresponding Units	\$400 million
100-Percent Follow-up on Vacant Housing Units	\$200 million
100-Percent Follow-up on Incomplete Questionnaires for Coverage	\$150 million
Expand Partnership Activities	\$25-50 million
Expand Promotion Activities	\$50-100 million
Deploy Special Enumeration Activities	\$25-50 million
Greater Quality Assurance	\$25-50 million
Eliminate Integrated Coverage Measurement (ICM)	- \$325 million
Conduct 1990-style Postcensus Evaluation Study	\$125 million
<b>Total:</b>	<b>\$675-800 million</b>

## **VII. AN ILLUSTRATION: THE MILWAUKEE COMPLETE COUNT CAMPAIGN**

The Milwaukee Complete Count Campaign illustrates the benefits and limitations of increased public outreach efforts in support of the census. The city government of Milwaukee, Wisconsin, was one of the Census Bureau's most enthusiastic supporters for the 1990 Census. Its "Complete Count" outreach campaign began in October, 1989, with a rally, peaked with a wide variety of activities from February through April, 1990, and ended with a "Were You Counted?" campaign in June, 1990. The cost to the city was about \$300,000, plus another \$55,000 of in-kind services. Additional donated and discounted services were valued at \$62,000.

While Milwaukee's efforts increased its count, they also demonstrated that the best of efforts still leave a sizable undercount. Even with Milwaukee's unprecedented efforts, the Post Enumeration Survey indicated that approximately 2.3 percent of the city's residents were missed, which was higher than the national average.

Research on 1990 Census outreach efforts such as those in Milwaukee has found that such efforts increased mail response, but did not eliminate coverage error. Milwaukee's efforts probably raised the mail return rate because its return rate of 76 percent was slightly higher than the national average of 74 percent (among the 26 cities of 500,000 or more, Milwaukee ranked third in mail return rates). A higher mail return rate still left 24 percent of occupied units in the city to be counted by door-to-door enumerators, and not every Milwaukee census tract had 24 percent nonresponse to mail questionnaires. Those tracts with minority populations of 80 percent or more had a mail nonresponse rate of 40 percent, almost four times the 11 percent rate in Milwaukee tracts with less than 10 percent minority populations.<sup>4</sup>

Administrative data also confirm a bias with traditional census methods, even with an aggressive outreach campaign such as that conducted by Milwaukee in 1990. An academic study compared the number of children found in Milwaukee in the 1990 Census with the number of children on Aid to Families with Dependent Children records. The study found that the 1990 Census undercounted children in 52 selected census tracts.

Was Milwaukee's Complete Count Campaign a success? To the extent that the campaign boosted mail response, it was a success. But the Milwaukee campaign also demonstrated the limits of spending money on outreach. Outreach has the potential to improve accuracy by boosting the mail return rate; it does not eliminate the undercount or the differential undercount.

---

<sup>4</sup> The mail return rate is calculated by dividing the number of mail returns by the number of housing units that were occupied. This rate differs from the "Mail response" rate mentioned earlier in this report that is calculated by dividing the mail returns by all housing units (including vacant units and deleted units). The mail response rate is important to use during census operations to determine quickly the nonresponse follow-up workloads.



## VIII. EXPECTED ERROR RATES

Errors in the census can arise from many sources—from respondents misunderstanding instructions, declining to participate, or giving inaccurate answers; from enumerators, postal workers, telephone operators, and data processors who make mistakes; from incorrect address lists; from poorly-worded questions; and from the way a census is planned and implemented.

Errors are grouped into two basic types -- those that occur during the measuring or data collection process (**nonsampling error**), and errors that occur because only part of the population is being directly contacted (**sampling error**). Nonsampling error occurs in both censuses and sample surveys; sampling error only occurs in sample surveys. Nonsampling errors can be the most serious types of errors because they yield biased results when most of the errors distort the results in the same direction. Decennial censuses have traditionally experienced nonsampling errors, most notably coverage error, or undercount, resulting from persons missed or double counted during the enumeration process.

### A. Nonsampling Error

Reducing and measuring nonsampling error is more complex than measuring and reducing sampling error. Sampling error can be measured, which is the reason it often gets more attention than nonsampling error. But nonsampling error and its consequent biases are present throughout the census process and can reduce the quality of results more than sampling error.

Sources of nonsampling error include:

1. **Coverage Error.** The 1990 Census and earlier censuses have been criticized for coverage error, that is, for missing people and housing units. This coverage error arises from two general problems -- missing entire housing units and missing some or all of the people in an enumerated unit. Based on the 1990 PES results, 69.5 percent of the coverage error came from enumerated housing units and the remaining 30.5 percent came from housing units that were not enumerated at all.

2. **Nonresponse Error.** Nonresponse error occurs when (1) housing units or people cannot be located or refuse to participate, or (2) answers to one or more items on the questionnaire are missing.

3. **Observational Error.** Observational error occurs when the questions asked on a census or survey yield inaccurate answers. These kinds of errors can be attributed to the interviewer, the questionnaire, the respondent, or the means by which the data are collected (telephone, personal visit, mail).

**4. Data Processing Error.** Data Processing error occurs after the data are collected, as a result of actions of processors -- during data entry, coding, editing, tabulation and other processing activities. Errors can also be introduced when missing data items are created from statistical modeling procedures (*i.e.*, imputation).

## **B. Sampling Error**

With any sample, scientifically selected or not, differences are likely to exist between the characteristics of the sampled population and the larger group from which the sample was chosen. However, in a scientific sample, sampling error is readily measured based on the mathematics of probability. Estimates of sampling error are referred to as sampling variance and are commonly expressed as the standard error as a percent of the population total. To a certain extent, sampling error can be controlled -- samples can be designed to ensure comparable levels of error across groups or across geographic areas.

## **C. Error Related to Estimation**

**1. Model Error.** Model error results from the use of statistical techniques to apply what one has learned from a sample of the population to improve the numbers for the entire population. For Census 2000, this type of error can arise in using ICM results to improve the census totals. The effect of model error is that the improvements being made to the totals are not perfect. The accuracy of ICM is based on the assumption that all individuals have the same chance to be included in the initial collection phase or in the ICM. The chances of inclusion can be different for the two systems. Because past experience with differential undercount demonstrates that some groups in the population are more likely to be missed in any physical enumeration, the Census Bureau defines poststrata. Poststrata are groupings based on variables that previous studies have shown to be related to coverage error. Examples of poststrata are renter-owner status, race and Hispanic origin, age, sex and urban-rural residence. Dual System Estimation assumes that the probability of being included is uniform within these poststrata. However, there still are differences among the individuals grouped within each poststratum. These differences are called "heterogeneity."

Heterogeneity within a poststratum affects the census totals in two ways and can result in an overestimate or an underestimate of the count. First, the estimation within a poststratum can fail to capture the variation in coverage error among small areas. Second, heterogeneity can make the poststratum estimate too low, so that not all misses are measured. This second effect is called correlation bias. Correlation bias is caused by the fact that people missed in the initial mail response and nonresponse follow-up are also more likely to be missed in the ICM survey. This problem leads to lower estimates of the undercount. Several research projects are in progress to assess and deal with error stemming from heterogeneity.

**2. Matching Error.** Matching error occurs, for example, when the ICM results are compared with the results of the initial phases of the enumeration. A person could be "in" both systems in reality, but only identified in one -- either the

initial phase or the ICM. When a difference is detected between the two data sets, follow-up interviews are conducted to resolve the inconsistencies. New computer technologies in Census 2000 for unduplication should reduce matching error.

**3. Contamination Error.** Contamination error occurs when there are two separate data collection activities. This error occurs when inclusion in one collection affects the response in the other collection.

#### **D. Gross Error Versus Net Error**

There are three types of coverage error -- omissions, duplicates, and erroneous inclusions. Omissions occur when housing units or people are missed. Duplicates occur when housing units or people are included more than once. Erroneous inclusions occur when people are incorrectly included in the initial enumeration because they are fictitious, in the wrong geographic location, etc. These types of errors can be combined to produce either net error or gross error numbers. Gross error refers to the total number of errors made in the census, while net error refers to the total effect of these errors on the resultant statistics. For gross error, the effect is additive; that is, the sum of people omitted plus duplicates plus erroneous inclusions. For net error, the errors are treated as an excess (duplicates and erroneous inclusions) or deficit (omissions), depending on the type of error, and the effect of combining produces a canceling-out effect. Gross error measures the total number of mistakes; net error measures the undercount.

The 1990 Post Enumeration Survey (PES) was designed to measure the net undercount in the 1990 Census by population group and to provide the data to adjust for that net undercount. However, due to interest in the level of gross errors, the data have been used to provide estimates of gross omissions, gross erroneous inclusions, and total gross coverage error. As the PES was not designed to estimate gross errors, there was no specific method for obtaining these estimates. Care must be taken in interpreting the gross error numbers; some of the measures and concepts are appropriate only when considered in terms of the way they produce net estimates, and all the PES numbers are subject to sampling error.

The Census Bureau cannot measure precisely and separately the effects of all the types of error described above. To the extent that it can measure and compare gross error, the Bureau has reviewed the gross (combined) error for 1990 and estimated the likely net error for 2000 with and without sampling.

The 1990 Census had a net undercount of approximately 4 million people. This figure is called "net" because it is the difference between the number of residents who were not counted at the geographic location being considered, less those residents incorrectly included or counted twice at the geographic location being considered. The gross error in the 1990 Census, however, was more than 26 million people: 15 million people were not counted at all, or were not counted in the correct block, while 11 million people were incorrectly included in a block. The incorrect inclusions may have been counted in more than one block or merely

assigned to an incorrect block. The net number of people not included in the national total represents the 4 million national net undercount.

Many users of census data are interested in gross error based on larger geographic areas, that is, many users do not care whether the people are counted in the correct block. When one ignores errors associated with individual blocks or other small areas, the gross error in 1990 was about 12.8 million people, with 8.4 million people not counted and 4.4 million people counted twice or incorrectly included in the census.

Looking to Census 2000, the Census Bureau has estimated the likely net error (from those types of error that can be measured) both for its plan and for a physical enumeration plan. A plan without sampling would include all the modernization plans for Census 2000 except for those that involve sampling. A modern non-sampling plan would cost \$675 to \$800 million more than the current sampling plan (additional costs detailed in Section VI(D)) and would be substantially less accurate.

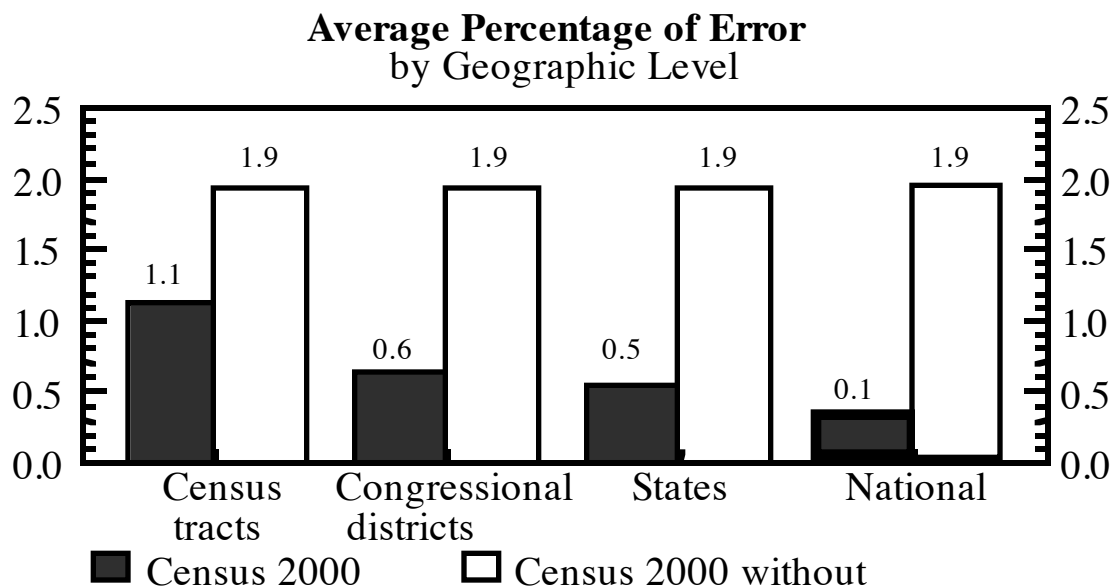
#### **E. Summary of Estimated Error, by Geographic Level Down to the Census Tract, for Plan Alternatives**

The Census Bureau and the Academy believe that the introduction of a limited use of sampling will make the census more accurate at the geographic levels for which its data are most critical: national, state, and Congressional district. With sampling, the estimate is just as likely to be above the true population as below it. The error declines as small areas are added together to create larger ones, such as Congressional districts.

All error figures in the table below were derived using simulations of 1990 Census estimates of undercounts and overcounts for census tracts. The Bureau has concluded that the error from the proposed plan will be 1.1 percent at the census tract level, 0.6 percent at the Congressional district level, 0.5 percent at the state level, and 0.1 percent at the national level. In contrast to sampling, physical enumeration methods are more likely to result in an underestimate of the population regardless of the size of the population area. The projected error from a physical enumeration and with no sampling in 2000 would average 1.9 percent at all levels from the census tract level to the national level.

	National	States	Congressional Districts	Census Tracts*
The Census 2000 Plan	0.1 %	0.5% (0.2% - 0.5%)	0.6% (0.3% - 2.3%)	1.1% (0.6% - 2.4%)
Improved procedures without any sampling	1.9 %	1.9% (0.4% - 3.2%)	1.9% (-1.2% - 7.0%)	1.9% (-1.2% - 6.2%)

\*The range of error at the census tract level has been "trimmed" so that it does not include the most extreme outliers--the highest and lowest 3 percent.



Note that this simulation exercise does not produce the same errors for all states, Congressional districts, or tracts. The estimates shown in the table are the average errors at each level. The numbers in parentheses show a high and low range of estimated errors at each level on a consistent basis. The “Combined Error” figures for Congressional districts and census tracts do not include error due to modeling; model error would not apply to the nonsampling alternative.

To account for expected growth in the population of the United States through the year 2000, the 1990 census tract population totals by race and Hispanic origin were projected using factors derived from a widely used process known as Demographic Analysis. The simulations assume that the percentage undercounts (and overcounts) measured for each group in the 1990 PES also would apply in Census 2000. To determine the amount of undercount or overcount for each census tract, the projected population totals for each were computed with and without the results of the 1990 PES for each region and for the various segments of the population within it. The totals for the specific census tracts in each geographic entity were summed to derive the error rates for more populous geographic levels, such as Congressional districts, states, and the nation.

In addition to foregoing improvements in accuracy, a ban on statistical procedures in Census 2000 would have other wide-ranging effects. Such a ban would preclude the cost-effective use of statistical sampling to check the vacancy information provided by the USPS. It would preclude the two most significant operations planned to reduce costs and improve the accuracy of Census 2000: the use of statistical sampling to finish the initial task of making contact with someone at an address; and the use of statistical sampling to account correctly for those individuals who are missed or counted more than once during the initial operation and follow-up. The bottom line is

that a ban on scientific sampling for Census 2000 would make the census less accurate than it could be, and more costly than it should be.

#### **F. Error Rates at the Block Level**

Given the constitutional purpose of the census to serve as the basis for apportioning the 435 seats in the House of Representatives among the states, the census is designed to maximize accuracy at the state level. Because census results are also important for drawing state legislative district boundaries and allocating grants to substate jurisdictions, the Census Bureau must also concern itself with accuracy for smaller geographic areas.

Neither the traditional methods used in the 1990 Census nor the scientific sampling methods planned for 2000 have emphasized accuracy at the block level. This lack of emphasis is appropriate because the population of stand-alone blocks is not used to determine legislative districts or to distribute population-based funding.

Error rates were quite substantial at the block level in the 1990 Census. The blocks in the Post Enumeration Survey had an average error of 7.6 percent. In many cases, a housing unit was assigned to the wrong block, which contributed to error in two blocks, one positive and the other negative. Just as with sampling error, large block level errors due to assignment to the wrong block tend to cancel each other out when blocks are aggregated together. When blocks are aggregated to census tracts and larger geographic areas, sampling errors decline sharply. The percentage error rate also falls when aggregating from the block level to the tract level using only traditional methods, but it does not decline as sharply as sampling error does, so that a sizable undercount remains. Thus, at the census tract level and larger areas of geography, the Census 2000 plan is more accurate than a census without sampling.

The Academy Panel on Alternative Methodologies warned against putting too much emphasis on accuracy at the block level in its Second Interim Report last month:

The important point to note here is that for the counts for census blocks, the level of sampling error is, relatively speaking, not an appropriate criterion for judging the quality of the census. Although block counts may contribute to the congressional redistricting process, for example, it is important to keep in mind that the results in a redistricting process are the counts for the congressional districts that are eventually created (and to a lesser extent, the counts for districts that were, or conceptually might have been, considered but were discarded). For these kinds of counts, the level of sampling error will be modest because the larger the number of observations used for an estimate, the smaller its sampling error will be.

Thus, in the panel's view, the important considerations for evaluating whether the amount of sampling error present in the census process is acceptable are not those that relate to counts for very small units, such as

blocks. It is clear that at that level, sampling error may be substantial in some cases (again, relative to the size of the block). The evaluation of sampling error should take place for the geographic level counts that have important legal, political, or financial implications. For such levels, a census that uses sampling can achieve results that are at least as good as those from a more time-consuming and expensive effort to obtain a completed form for every household.

Questions have been raised about calculated error rates at the block level in the 1995 Census Test of some census methods. The 1995 Census Test, conducted in Oakland, California, Paterson, New Jersey, and six parishes in Northwest Louisiana, was designed as an experiment to test various methodologies for sampling for nonresponse follow-up and Integrated Coverage Measurement; the test was not an attempt to demonstrate Census 2000 sampling methods. In particular, the 1995 test has no relevance for the important question of whether a census based on scientific sampling is more accurate than a census based only on physical enumeration. The only estimates of error at the block level in the 1995 test result from statistical theory based on the number of households in the sample for each block. The estimates do not compare the results of sampling with the true population total at the block level, an unknown number.

The following table shows the block level sampling error of the samples used in the 1995 Census Test:

<b>Summary Statistics for the Block Level Coefficients of Variation 1995 Census Test Sites</b>			
Test Site	Weighted Average	Minimum	Maximum
Paterson	0.1828	0.0559	2.4336
Oakland	0.1262	0.0403	1.4359
NW Louisiana	0.2519	0.0438	1.5142

Source: 1995 Census Test

The nonresponse sample sizes in each site for the 1995 Census Test were much smaller than the Census 2000 design would require. In Oakland and Paterson, the Census Bureau sampled about one in three nonresponding housing units, while in Northwest Louisiana, it sampled one in four. Under the Census 2000 design, the Census Bureau would have sampled approximately three out of every four nonresponding housing units in Oakland and Northwest Louisiana; the rate in Paterson would have been five in six, given initial response rates in these three sites.

The implication of the smaller sample sizes is that the block level error rate from the 1995 Census Test is substantially greater than the expected error rate using the Census 2000 design. The effects are illustrated by comparing what was achieved in these three sites to the sampling error expected in Census 2000. For Paterson and Northwest Louisiana, the 1995 errors were three times or more larger than anticipated 2000 errors. For Oakland, the errors in 1995 were about

2.5 times as large. These calculations are based on the current Census 2000 design of reaching 90 percent direct response in each census tract.

Calculations of average error for small blocks can be misleading to those unfamiliar with these statistics. Statistical theory expects that small samples from small population blocks will have large errors. But a large percentage error of a small population number is still a small number. For example, the block in Paterson, New Jersey, that had an estimated error rate of 243 percent had only one resident. Because small blocks contribute less in any aggregations, the large percentage errors for some small blocks have little practical effect.



## **IX. PROCEDURES TO ENSURE UNBIASED STATISTICAL DECISIONS**

Concern has been expressed that the formulas for drawing samples or for extrapolating from sample data in the Census 2000 could be surreptitiously manipulated for political ends.<sup>5</sup> This concern is misplaced. Every effort has been made to ensure the independence and integrity of the decisions by the professional statisticians at the Census Bureau. The professionals have been -- and will remain -- insulated from political interference throughout the Census 2000 process.

The Census Bureau has a long history of political independence. The Bureau refused, when requested during World War II, to identify Japanese-American individuals from Census Bureau records, relying on its specific confidentiality requirement in Title 13 of the U.S. Code. In the past, the Bureau resisted attempts to manipulate poverty statistics. The Census Bureau is staffed by many of the world's preeminent professional statisticians and demographers; it is a professional organization with a long history of scientific integrity. It is worth noting also that none of the myriad lawsuits brought after the last three decennial censuses resulted in the Census Bureau changing its final enumeration. The Census Bureau will continue to resist any attempts to manipulate its data or processes; accuracy will guide all operational decisions in Census 2000.

In fact, experts agree that the use of sampling in Census 2000 should minimize the opportunity for political manipulation, not increase it. Sampling has known, objective properties. The known properties of sampling are preferable to the certainty of missing several million people using traditional counting methods alone. In fact, uncontrolled error is more of a concern with a traditional headcount than it is with sampling.

The basic statistical framework for Census 2000 was developed by the Census Bureau in tandem with the Academy. The Academy study mandated by Congress in 1991 recommended the major innovations for Census 2000. Concluding that continued reliance on traditional enumeration methods was futile, the Academy recommended that the Bureau adopt both sampling for nonresponse follow-up and

---

<sup>5</sup> The decennial census is of immense importance to political representation and federal funds distribution. The census' very importance virtually ensures that political groups will be vitally interested in the outcome of each decennial. The 1920 Census is a vivid illustration of the intense political interest in the census. The shift in the nation's population between 1910 and 1920 from rural areas and foreign lands to America's urban areas so disturbed the balance of power in Congress that the House of Representatives was never able to agree on a plan to reapportion itself. Congress essentially ignored the results of the 1920 Census and did not reapportion until after the 1930 Census. Congressional debate over the proper method of apportionment ultimately lead to request for a National Academy of Sciences report, the results of which became the basis of the reapportionment statute, 2 U.S.C. § 2a. Title 2 now provides for automatic reapportionment of the House of Representatives upon the reporting of the census numbers by the President to Congress.

sampling with a very intensive quality check process to provide a statistical enumeration more accurate than is possible with a physical enumeration. The Census Bureau has spent the last five years turning the Academy's initial and ongoing recommendations into a detailed plan. The Bureau has also been working with its advisory committees, conducting public meetings, and having discussions with Congress to refine the details of the plan.

The Census Bureau is committed to making its decisions on formulas in a very open process. This openness should prevent not only the possibility of surreptitious manipulation, but the perception of possible manipulation.

The Census Bureau has proposed to the National Academy of Science's Committee on National Statistics that it convene a fourth expert panel to guide the Bureau's work for Census 2000. This new group of outside experts would critically review the statistical procedures for the Census 2000, in particular the use of sampling for nonresponse follow-up and Integrated Coverage Measurement. These experts will comment on the planning process, suggest improvements and preferred approaches, and review other procedures that may be considered during the enumeration in order to increase accuracy. The panel will be established in the Fall of 1997 and continue its work through the Spring of 2001, with the reporting of census results for reapportionment and redistricting. This open, expert review will:

- assure the objectivity, scientific validity, and integrity of the 2000 Census,
- assist the Bureau in its goal of producing a more accurate Census, and
- improve understanding of how sampling and statistical estimation procedures contribute to achieving a more accurate Census.

The panel will consist of widely and highly regarded experts on census matters, statistical methodology, sampling, survey research, demography, and other social and behavioral sciences. The panel will interact with the Census Bureau as follows:

- Census Bureau will develop statistical procedures for Census 2000.
- The panel will convene periodic open workshops to review specific procedures, inviting other experts and various stakeholders to attend and critique the Bureau's procedures.
- Following each workshop, the Census Bureau will, as necessary, revise its planned procedures based on issues raised and suggestions made at the workshop and then resubmit the procedures to the panel.
- The panel will review the Census Bureau's revised procedures, other documentation, and the workshop proceedings and then issue its assessment of the specific procedures planned. These reports will be

reviewed by the Committee on National Statistics, the Commission on Behavioral and Social Sciences and Education, and other expert groups in accordance with National Research Council procedures.

- The Census Bureau will finalize its procedures prior to the Census, based on the recommendations of the panel.
- During the conduct of the actual enumeration, the panel will review the statistical procedures as they are implemented, as well as other procedures that may be considered in order to increase accuracy.

The Academy Panel on Alternative Methodologies concluded that this process of outside peer review would remove potential objections to the use of sampling:

If sound procedures are developed by the Census Bureau and communicated to users, the panel believes that it will be possible for the Bureau to address all reasonable potential objections to the uses of sampling and to satisfy users that the use of sampling has added to the soundness and quality of the 2000 Census, rather than detracting from it.

By mid-1998, the Bureau will make all its planned formulas available for scrutiny by the public, the professional statistical community, and the new Academy panel. The Bureau will consider all comments and criticisms for more than a year. Then, based on the best professional judgment at the time, the Census Bureau will announce and “lock in” its final set of formulas -- well in advance of the collection of any data in 2000. Fears that the Census Bureau will collect data in 2000 and then use new formulas designed to achieve some purpose other than the most accurate census possible are completely without foundation.

## **X. LEGAL CONSIDERATIONS**

The plan for Census 2000 is both Constitutional and legal.

The Assistant Attorney General for the Civil Division in the Bush Administration, Stuart Gerson, concluded in a July 9, 1991 opinion that the Constitution's requirement of an actual enumeration refers to the accuracy of the census, not to any particular method of census taking. In addition, the Assistant Attorney General concluded that the weight of caselaw on the Census Act does not prohibit adjustment. Mr. Gerson detailed his careful examination of the Constitution's requirement for an "actual enumeration" in testimony before the Senate Committee on Governmental Affairs on April 16, 1997. He explained that, at the time the Constitution was written, the term "actual" meant both "existing in act or fact" and "in action or existence at the time, present, current." He noted that Georgia was seeking representation in the Congress to be formed based on Georgia's expected population growth rather than its current population. He concluded that the term "actual" suggest the 'Framers' intent that the census be based on current population, as opposed to taking into account potential population growth. It does not appear to delimit the means by which an accounting of the currently existing population may be determined.

In evaluating the term "enumeration," Mr. Gerson similarly found no reason to favor sole reliance on physical enumeration to the exclusion of statistical sampling:

In sum, the essence of enumeration, as the term is both generally and constitutionally understood, is more likely found in the *accuracy* of census taking rather than in the selection of any particular method, i.e., a headcount. [emphasis added]

In 1994 the Department of Justice (DOJ) reviewed the Census Bureau's preliminary plans to use sampling in Census 2000 and issued a written opinion confirming that the plan was neither illegal nor unconstitutional. This DOJ opinion is premised on a long line of federal court cases holding that neither the Constitution nor the Census Act bars the use of sampling in a decennial census, so long as sampling is not used as a substitute for a traditional enumeration.

In addition, the Supreme Court recently resolved the most prominent case challenging the 1990 decennial Census, Wisconsin v. City of New York, 116 S. Ct. 1091 (1996). While the Court's opinion did not directly address the legality of sampling, the Court confirmed that the Secretary of Commerce enjoys broad discretion in the methods used to take the census. In the exercise of this discretion, the Census Bureau has determined that it cannot take the most accurate and cost effective Census possible without a limited, judicious use of sampling. The Bureau proposes to use sampling in Census 2000 as a complement to traditional methods used in enumeration, not as a substitute for these methods.

The Constitution requires that an "actual Enumeration" be conducted every ten years "... in such Manner as [the Congress] shall by Law direct." The actual enumeration requirement is not a requirement to conduct a headcount or physical enumeration. Courts that have considered this issue have unanimously concluded that actual enumeration means that the decennial census must be as accurate at that time as possible, without reference to the specific method that is used.<sup>6</sup>

The Census Act, Title 13 of the United States Code, is the statutory vehicle through which Congress delegated responsibility for conducting the Census to the Secretary of Commerce. Section 141(a) requires the Secretary to take a decennial Census "in such form and content as he may determine, including the use of sampling procedures and special surveys...", while Section 195 mandates that the Secretary "shall, if he considers it feasible, authorize the use of ... sampling ..." except "for the determination of population for purposes of apportionment of Representatives." Courts have held that these provisions, taken together, evidence Congress' intention that sampling may be used in a decennial Census so long as it is not a substitute for traditional methods of numeration.<sup>7</sup>

In 1994, DOJ specifically approved the Census Bureau's plan to use sampling in Census 2000, agreeing with the logic contained in the long line of court decisions holding that neither the Constitution nor the Census Act prohibits adjustment:

---

<sup>6</sup> Young v. Klutznick, 652 F.2d 617 at 625 (6th Cir. 1981) ("[A]lthough the Constitution prohibits subterfuge in adjustment of Census figures for purposes of redistricting, it does not constrain adjustment of Census figures if thoroughly documented and applied in a systematic manner."); City of New York v. U.S. Department of Commerce, 739 F. Supp. 761 at 767 (E.D.N.Y. 1990), rev'd, 34 F.3d 1114 (2nd Cir. 1994), 116 S. Ct. 1091 (1996), ("It is no longer novel, or in any sense new law to declare that statistical adjustment of the decennial Census is both legal and constitutional ...."); Carey v. Klutznick, 508 F. Supp. 404 at 415 (S.D.N.Y. 1980) ("It appears to the Court that [the Constitution's requirement for an actual enumeration] indicates an intent that apportionment be based on a Census that most accurately reflects the true population of each state."); City of Philadelphia v. Klutznick, 503 F. Supp. 663 at 679 (E.D. Penn. 1980) ("[I]t is inconceivable that the Constitution would require the continued use of a headcount in counting the population.").

<sup>7</sup> City of New York v. U.S. Dept. Of Commerce, 34 F.3d 1114 at 1125 (2nd Cir. 1994), rev'd on other grounds, 116 S. Ct. 1091 (1996) ("statistical adjustment to the initial enumeration is not barred by the Census Act and indeed was meant to be encouraged."); Carey v. Klutznick, 508 F. Supp. 404 at 415 (S.D.N.Y. 1980) ("the Census Bureau [is authorized by § 195 to] ... utilize sampling procedures but only in addition to more traditional methods of enumeration."); Young v. Klutznick, 497 F. Supp. 1318 at 1335 (E.D. Mich. 1980) rev'd on standing, 652 F.2d 617 (6th Cir. 1981) ("All that § 195 does is prohibit the use of figures derived solely by statistical techniques. It does not prohibit the use of statistics in addition to the more traditional measuring tools to arrive at a more accurate population count.").

[I]n requiring an 'actual' enumeration, the Framers meant a set of figures that was not a matter of conjecture and compromise.... There is no indication that the Framers insisted that Congress adopt a 'headcount' as the sole method for carrying out the enumeration, even if later refinements in the metric of populations would produce more accurate measures....

...[T]he Census Act does not preclude the Bureau from engaging in statistical adjustments of the next set of decennial Census figures.... Its prohibition on 'sampling' in decennial Censuses appears to have meant only that while a procedure relying on 'sampling' alone might be the most cost-effective means to discover the information sought in a mid-decade Census, the Bureau should not rely on 'sampling' as its exclusive method of tabulating population figures in the decennial Census....

This 1994 opinion is in accord with earlier DOJ opinions holding that the Department of Commerce could have adjusted the 1980 and the 1990 Censuses, had it determined that adjustment was feasible and proper.

Finally, some have contended that being counted in the census is like voting, and each individual can decide for him or herself whether to participate in the political process by being counted. This argument not only fails to recognize that children under the age of 18 accounted for more than half of the undercount in 1990, but also fails to comprehend the Census Bureau's constitutional mandate. Article I, Section II of the Constitution, as amended by the Fourteenth Amendment, commands that representatives be apportioned based on the "whole number of persons in each State", not the number of persons in each state who choose to participate in the political process. The apportionment situation at the time the Constitution was adopted makes clear that the Framers' intent was to count all living persons in the United States, not all voters or all citizens. At the time the Constitution was adopted, women were not permitted to vote; children and slaves were in no position to "stand up and be counted." All were nonetheless enumerated for apportionment purposes, with slaves being counted as 3/5 of a person until passage of the Fourteenth Amendment in 1868. The Constitutional command to the Census Bureau is clear -- to secure the most accurate enumeration possible of all persons regardless of status, so that Congress can reapportion itself fairly. The Census Bureau's goal is to find and enumerate all persons resident in the United States on Census Day, 2000.

The legal authorities are clear -- neither the Constitution nor the Census Act precludes the use of sampling. The Census Bureau, following the Congressionally-mandated recommendations of the National Academy of Sciences, has determined that Census 2000 would be rendered more accurate and more cost-effective by the introduction of a limited use of sampling in addition to a traditional methods of enumeration. The Supreme Court has held that the Census Bureau enjoys broad discretion in the methods it uses to take the census. The decision to use sampling as planned in Census 2000 is a rational decision and falls well within this discretion.

## **GLOSSARY OF TERMS**

### **Computer Assisted Telephone Interviewing (CATI)**

A method of data collection using telephone interviews in which the questions to be asked are displayed on a computer screen and responses are entered directly into the computer.

### **Data Access and Dissemination System (DADS)**

A generalized electronic system for all access and dissemination of Census Bureau data. This interactive electronic system will be designed to allow efficient and cost-effective access to data summaries generated by the various censuses and other programs of the Census Bureau. DADS will serve as the vehicle for accessing and disseminating data from Census 2000 and from the American Community Survey.

### **Demographic Analysis (DA)**

Demographic Analysis is one of the methods the Census Bureau uses to measure coverage at the national level. It differs from survey coverage estimates, such as PES and ICM, in that it does not rely on case by case matching of census records. To produce an estimate of the total population, DA relies on administrative records to provide estimates of births, deaths, immigration, and emigration. DA provides estimates on the national level only.

### **Dual System Estimation (DSE)**

The estimation methodology used for Integrated Coverage Measurement.

### **Geocoding**

The assignment of an address, structure, key geographic location, or business name to a location that is identified by one or more geographic codes.

### **Group quarters**

A facility where people live that is not a typical household-type living arrangement. The Census Bureau classifies all individuals not living in households as living in group quarters. There are two types of group quarters: institutional (for example, correctional facilities, nursing homes, and mental hospitals) and noninstitutional (for example, college dormitories, military bases and ships, hotels, motels, rooming houses, group homes, missions, shelters, and flophouses).

## **Heterogeneity**

Heterogeneity occurs when blocks of housing units assigned to sampling strata or groupings are not similar in terms of the likelihood of being included or missed by the census. Heterogeneity creates difficulty for the small area estimation process because the correction factor gets applied to all people with the specified characteristic in that sampling poststratum, even though some of them do not actually have the coverage characteristics.

## **Homogeneity**

The assumption of homogeneity expects that all people in a particular sampling stratum or grouping will be very much alike in terms of their likelihood of being included or missed by the census. The grouping of people in a particular stratum is called poststratum, such as all white, non-Hispanic male renters ages 18-22 in a rural area. A lack of homogeneity in a particular sample block is not an error, but it does create difficulty for the small area estimation process. This happens because the correction factor gets applied to all people with the specified characteristic in that poststratum, even though some of them do not exhibit the same coverage characteristics.

## **Housing unit**

A housing unit is a house, an apartment, a mobile home or trailer, a group of rooms, or a single room occupied as a separate living quarters, or if vacant, intended for occupancy as a separate living quarters. Separate living quarters are those in which the occupants live separately from any other individuals in the building and which have direct access from outside the building or through a common hall. For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible.

## **Imputation**

When information is missing or inconsistent, the Census Bureau uses a method called imputation to assign values. Imputation relies on the statistical principle of “homogeneity,” or the tendency of households within a small geographic area to be similar in most characteristics. For example, the value of “rented” is likely to be imputed for a housing unit not reporting on owner/renter status in a neighborhood with multi-units or apartments where other respondents reported “rented” on the census questionnaire. In past censuses, when the occupancy status or the number of residents was not known for a housing unit, this information was imputed.

## **Integrated Coverage Measurement (ICM)**

A coverage measurement methodology that will be used to determine the number of people and housing units missed or counted more than once in Census 2000. This information is combined with the initial data collection results before producing a single set of official census results (the one-number census).



**List/enumerate**

A method of data collection in which temporary field staff, called enumerators, list each residential address, spot the location of each on a census map, and interview the residents of the household during a single visit. This completes the census address list for these areas and provides the information needed to update the TIGER data base and Master Address File (see definitions below).

**Local Update of Census Addresses (LUCA)**

A Census 2000 program, established in response to requirements of P.L. 103-430, that provides an opportunity for state, local, and tribal governments to review and update individual address information in the MAF and associated geographic information in the TIGER data base before using the addresses for questionnaire delivery to improve the completeness and accuracy of both computer files and the census.

**Master Address File (MAF)**

A computer file based on a combination of the addresses in the 1990 census address file and current versions, supplemented by address information provided by state, local, and tribal governments. The MAF is being updated throughout this decade and the next to provide a basis for creating the Census 2000 address list, the address list for the American Community Survey, and the address list for the Census Bureau's other demographic surveys.

**Nonresponse follow-up**

A census follow-up operation in which temporary field staff, known as enumerators, visit addresses from which no response was received.

**Nonsampling error**

Errors that occur during the measuring or data collection process. Nonsampling errors can be the most serious types of errors because they yield biased results when most of the errors distort the results in the same direction. Unfortunately, the full extent of nonsampling error is unknown. Decennial censuses traditionally have experienced nonsampling errors, most notably undercount, resulting from people being missed in the enumeration processes.

**Post-Enumeration Survey (PES)**

The 1990 Post-Enumeration Survey (PES) was designed to measure net coverage errors in the 1990 census. The PES evaluated coverage in the 1990 census on a case-by-case basis using the Dual System Estimation (DSE) methodology.

## **Poststratum**

Information about the current occupants of each housing unit in the ICM survey found during the ICM interview, is used to form groupings called “poststrata.” This information, including the age of respondent, current owner/renter status, and so forth, is used to form homogenous groupings and improve the estimation process. By contrast, the initial ICM strata will be formed using aggregate information about each block as of the 1990 census.

## **Program for Address List Supplementation (PALS)**

A program providing all governmental units and regional and metropolitan agencies the opportunity to submit lists of individual addresses for their community to the Census Bureau for use in building the MAF. Ongoing submissions and feedback between the Census Bureau and local governments on this program, enabled by the Census Address List Improvement Act of 1994 (P.L. 103-430) will help ensure the completeness and accuracy of the MAF and the TIGER data base.

## **Quality assurance (QA)**

Quality assurance represents a broad philosophy and specific procedures that are designed to: build quality into the system, constantly improve the system, integrate responsibility for quality with production.

## **Sampling error**

Errors that occur because only part of the population is being contacted directly. With any sample, differences are likely to exist between the characteristics of the sampled population and the larger group from which the sample was chosen. However, sampling error, unlike nonsampling error, is readily measured.

## **Service-based enumeration (SBE)**

An operation designed to enumerate people at facilities where they might receive services, such as shelters, soup kitchens, health-care facilities and other selected locations. This operation targets the types of services that primarily serve people who have no usual residence.

## **Special place**

An institution that includes facilities where people live or stay other than the usual house, apartment, or mobile home. Examples are colleges and universities, nursing homes, hospitals, and prisons. Often the facilities that house people are group quarters, but they may include standard houses or apartments as well.

## **Sampling stratum**

A sampling stratum, as used in the ICM, is a grouping or classification that have a similar set of characteristics, based on the 1990 census. For example, one might

define a stratum as: all blocks in large central cities with a 1990 census population that was 30 percent or more Black renters.

### **Topologically Integrated Geographic Encoding and Referencing (TIGER)**

A computer data base that contains a digital representation of all census-required map features (streets, roads, rivers, railroads, lakes, and so forth), the related attributes for each (street names, address ranges, etc.), and the geographic identification codes for all entities used by the Census Bureau to tabulate data for the United States, Puerto Rico, and the Island Areas. The TIGER data base records the interrelationships among these features, attributes, and geographic codes and provides a resource for the production of maps, entity headers for data tabulations, and automated assignment of addresses to a geographic location in a process known as "geocoding."

### **Telephone Questionnaire Assistance (TQA)**

A toll-free service that will be provided by a commercial phone center to answer questions about Census 2000 and the Census 2000 questionnaire, and to take interviews from people who prefer to be interviewed over the telephone.

If you want a copy of any of the following reference materials, please call the Census Bureau at (301) 457-2131.

Modernizing the U.S. Census, National Research Council, 1995

Preparing for the 2000 Census, Interim Report II, National Research Council, Committee on National Statistics, 1997

Census 2000 Operational Plan, U.S. Bureau of the Census, 1997